Author:
Lowndes, David L D

Title:
Low cost, short range free space quantum cryptography for consumer applications

pocket size for pocket change

Low Cost, Short Range Free Space Quantum Cryptography for Consumer Applications: Pocket Size for Pocket Change

D.L.D Lowndes

October 10, 2014

Word Count: ≈ 35000

A dissertation submitted to the University of Bristol in accordance with the requirements for the degree of Doctor of Philosophy in the Faculty of Engineering, Department of Electrical and Electronic Engineering.

"History has taught us: never underestimate the amount of money, time, and effort someone will expend to thwart a security system. It's always better to assume the worst. Assume your adversaries are better than they are. Assume science and technology will soon be able to do things they cannot yet. Give yourself a margin for error. Give yourself more security than you need today. When the unexpected happens, you'll be glad you did."

Bruce Schneier.

Declaration and Copyright

Declaration

Unless otherwise acknowledged, the content of this thesis is the original and sole work of the author. No portion of the work in this thesis has been submitted by the author in support of an application for any other degree or qualification, at this or any other university or institute of learning. The views expressed in this thesis are those of the author, and not necessarily those of the University of Bristol.

David Leonard Dennis Lowndes

Abstract

A Quantum Key Distribution system has been demonstrated with a focus on applicability of the technology to a consumer use model. The optical devices were split into a large, expensive "Quantum ATM" (the Bob terminal) and a small, cheap handheld section (the Alice device).

Work was initially done to integrate the devices into a realistic demonstration system which was exhibited at several international conferences. Following this success, the devices constituent parts were isolated and improved upon. Work on the Alice device focused upon proposing a new scheme for light collimation and then investigating a method to enact this. In the Bob device the progress was directed towards a more effective single photon detection system utilizing active quenching.

An asymptotic secure key rate of 20kb/s was obtained at a bit error rate (BER) of 4% corresponding to a daylight operation scenario. A docking scheme was enacted to simplify the alignment of the optical channel which was shown to be repeatable over at least 50 cycles. The system was also shown to be stable over an extended period of time.

Contents

List of Figures

List of Tables

List of Abbreviations

APD avalanche photodiode

ATM automated teller machine

BB84 Bennett Brassard 1984

B92 Bennett 1992

BER bit error rate

BQP bounded error quantum polynomial time

CAP chip authentication program

CPLD complex programmable logic device

DOE diffractive optical element

ECL emitter-coupled logic

Eve eavesdropper

EMV Europay, MasterCard and VISA

FIB focussed ion beam

FPGA field programmable gate array

IC integrated circuit

LED light emitting diode

MITM man in the middle

ND neutral density

NFC near field communication

OCXO oven-controlled crystal oscillator

OTP one time pad

PBS polarizing beam splitter

PCB printed circuit board

PDA personal digital assistant

PER polarization extinction ratio

PKC public key cryptography

PNS photon number splitting

POS point of sale

QBER quantum bit error rate

QKD quantum key distribution

QRNG quantum random number generator

qubit quantum bit

RAM random access memory

SECOQC SECure COmmunication based on Quantum Cryptography

SEM scanning electron microscope

SMF single mode fiber

SPAD single photon avalanche diode

SSH secure shell

TCSPC time correlated single photon counting

TIA time interval analyser

TTL transistor-transistor logic

VNC virtual network computing

WCP weak coherent pulse

WGP wire grid polarizer

List of Publications and Presentations

29th June - 3rd July 2008: QIPIRC Summer School 2008, Sheringham, *Poster Title: "Low-cost devices for quantum cryptography"*

3rd - 4th July 2008: QIPIRC Conference 2008, Oxford, *Poster Title: "Low-cost devices for quantum cryptography"*

8th - 10th October 2008: SECOQC International Conference, Vienna, *Practical Demonstration*

12th - 15th January 2009: Rank Prize Funds Mini-symposium on Single Photon Detectors: Physics and Applications, Grasmere, *Talk Title: "A single photon detection system for application in a low cost quantum key distribution device"*

21st - 23rd April 2009: European Future Technologies Conference, Prague, *Exhibit Title: "Qubit Applications"*

6th - 7th April 2010: FIB 4 Photonics, Cambridge, *Talk Title: "Design, modelling and fabrication of sub wavelength pitch nanowire grid polarizers for application in compact quantum key distribution system"*

In preparation: Paper: "Low cost and short range Quantum Key Distribution"

Chapter 1

Introduction

1.1 Background

Most modern cryptography relies on the complexity of various mathematical operations to ensure security [1]. Presently this is an acceptable situation however recent advances in the field of quantum information have uncovered that were a quantum computer constructed, there are efficient algorithms [2] to solve the complex equations required to crack the current cryptosystems.

It is fortunate, however, that other advances in quantum information have provided an alternate scheme for cryptography, in which unconditional security is ensured by the laws of physics. Quantum key distribution (QKD) is the process by which a random bit string (key) is established between two parties, hereby known as "Alice" and "Bob", secure in the knowledge that this sharing cannot be intercepted by an eavesdropper, hereby known as "Eve".

There has been work on "post-quantum" cryptography which consists of cryptography based on problems whose solutions do not benefit from a speed increase on a quantum computer however since classical systems have been beaten once, it is conceivable that further advances in quantum computing could expose flaws in any

.	.
1	It must be small and light; comparable to current consumer cryptographic devices.
2	It must be low cost such that, for example, it could be distributed, for free, to customers of a bank.
3	Key agreement between the user and central terminal must be simple to initiate and quick to conclude.
4	These criteria must not compromise the unconditional security offered by QKD.

Table 1.1: The criteria for the design brief for our QKD system.

system implemented that is currently assumed to be secure.

In light of this, extensive and diverse research is being carried out worldwide to prepare a wide range of QKD solutions to enable widespread adoption of this unconditionally secure cryptographic technique. The work described herein specifically targets the requirement for low cost and compact devices suitable for the general public to carry out their daily cryptographic activities (such as banking or using the internet).

The requirements for such a system are rather different than those for a QKD system for, say, encrypting messages between government and military facilities; a list of the requirements used to design the system is presented in table 1.1.

1.2 Cryptography

History It is conceivable that since the invention of writing, people have devised methods with which to restrict who can read messages. Two practices related to this are *Steganography*, the art of hiding the existence of messages [3] and *Cryptography* which concerns itself with obscuring the meaning of a message [4].

To illustrate the difference, two examples from early history will be considered. An example of *Steganography* from ancient Greek history is the story of Histiaeus tattooing a message on a slave's head and sending the slave to Aristagoras once the hair had re-grown [5], this method masks the fact that a message has been sent at all however the message is freely readable to anyone who knows of its presence.

Cryptography, however simply renders a message gibberish to all but those who know how to decode it. One of the earliest examples, the *Caesar Cipher* is described later in section 1.2.1. In short this performs a simple substitution wherein letters are replaced with prearranged other letters. This message can then be sent openly as it can only be decoded by its intended recipient.

One Directional Problems Whereas a "symmetric" scheme utilizes the same key to encode and decode messages, there also exists an *asymmetric* method wherein the encryption key is not the same as the decryption key and the message cannot be decoded using only the encryption key [6]. The difference between these can be visualised by considering sending messages in locked boxes.

In the symmetric scheme, both the sender and receiver simply possess a key which opens a lock on a sealed box. The sender can then lock the box and upon receipt of the locked box the receiver can use their copy of the key to unlock it. The main issue of this system is the fact that the two parties need to have previously met in order to possess the same keys.

If we continue with the analogy, the asymmetric scheme can avoid this issue by starting with the receiver sending a customised open padlock to which they, and they alone own the key, to the sender. The sender can then lock the box using this lock and send the box safe in the knowledge that no parties other than the receiver are able to unlock it.

Knowledge and Computing power The holy grail of cryptography is the notion of "unconditional security" which is defined as a ciphertext which is secure even if the adversary has unlimited computing power [7]. Clearly, substitution ciphers based on some rule known to both parties can be broken by either a brute-force method of trying all possible substitutions or some sort of frequency analysis where the relative occurrences of certain letters in the ciphertext are compared with the relative occurrence of those letters in language.

This chapter discusses in detail the points discussed above and elucidates the approaches developed towards the goal of unconditional security and the place that this thesis holds within the environment of cryptographic applications.

1.2.1 Symmetric Schemes

1.2.1.1 Caesar Cipher

Symmetric cryptosystems are those in which the encryption and decryption keys are the same [6]. An early example of such is called the *Caesar Cipher* due to the stories surrounding Julius Caesar wishing to keep unwanted parties from reading battle orders sent over long distances. The process of the Caesar cipher takes each letter in a plaintext string and replaces it with a letter a fixed distance further along the alphabet [4] (see figure 1.1). The obvious disadvantage of this system is that once one knows that the Caesar cipher is being implemented, there are only 26 possible shifts to analyse for an attacker making it rather simple to break.

1.2.1.2 Vigenère Shift Cipher

Using the same idea as the Caesar cipher, the *Vigenère shift cipher* adds a further level of complexity by adding a different shift to each letter in the plain text corresponding to the letters in a keyword [6]. The keyword is repeated until it matches

Plaintext:	K	I	L	L	T	H	E	P	R	E	S	I	D	E	N	T
Convert letters to numbers, A=0, B=1,....Z=25.																
Plain Number values:	10	8	11	11	19	7	4	15	17	4	18	8	3	4	13	19
Number values +1:	11	9	12	12	20	8	5	16	18	5	19	9	4	5	14	20
Convert numbers back to letters.																
Ciphertext:	L	J	M	M	U	I	F	Q	S	F	T	J	E	F	O	U

Figure 1.1: Example of a Caesar cipher plaintext and ciphertext. The shift applied to the message is 1

the length of the plaintext and then each letter of the plaintext is shifted by the alphabetical value of the keyword text ($A = 1$, $B = 2$ etc.) A table of the values can be drawn to speed up manual encryption and decryption process (figure 1.2), encryption is performed by finding the plaintext letter in the first column and the keyword text letter in the first row and looking up the corresponding ciphertext value and decryption by the reverse. This cipher adds a deal of complexity over the Caesar cipher however if the key length is known, the problem can be broken down into a series of Caesar ciphers which can be analysed individually.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
G	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

Figure 1.2: The Vigenère shift cipher matrix renders encoding/decoding by hand much quicker; the effect of each key letter is mapped for each plaintext letter.

Plaintext:	K	I	L	L	T	H	E	P	R	E	S	I	D	E	N	T
Plain Number values:	10	8	11	11	19	7	4	15	17	4	18	8	3	4	13	19
Repeat encoding word:	S	E	C	R	E	T	S	E	C	R	E	T	S	E	C	R
Encoding number values:	18	4	2	17	4	19	18	4	2	17	4	19	18	4	2	17
Add number values mod 26:	2	12	14	2	23	0	22	19	19	21	22	1	21	8	15	10
Ciphertext	C	M	N	C	X	A	W	T	T	V	W	B	V	I	P	K

Figure 1.3: Example of a Vigenère cipher plaintext and ciphertext. The word used to encode the message is "SECRET"

1.2.1.3 One Time Pad

A Vigenère Shift cipher with a random key with length equal to that of the plaintext would be theoretically unbreakable however multiple reuses of the same key would allow for analysis by collecting several ciphertexts and analysing in the same way as previously. If the key were to never be reused again, this would constitute a *one time pad* system [4] (sometimes referred to as a *Vernam cipher* [8]), called so as one could imagine two parties sharing a note*pad* filled with pages of key where each page would be used *one time* and discarded. This would be 100% secure as there is no periodicity to analyze in the ciphertext.

1.2.2 Asymmetric Schemes

1.2.2.1 Diffie Hellman

Diffie-Hellman cryptography is based on the *discrete logarithm problem* [9], put simply, it relies on the fact that given "$g^a \pmod p$" and "$g^b \pmod p$" it is difficult to find "$g^{ab} \pmod p$" (terms defined below)

The secret, S, that is generated is possessed by Alice and Bob is a symmetric key however it is discussed in the section addressing asymmetric systems since the technique to obtain it is an asymmetric one. Once the method of obtaining S is explained, a Public Key (figure 1.4) implementation of Diffie Hellman will also be described. [1]:

Alice and Bob publicly choose p, g where p is prime, g is primitive root $\pmod p$ [1] [10]. They then secretly choose a, b respectively. They then calculate

$$A = g^a \bmod p \tag{1.1}$$

$$B = g^b \bmod p \tag{1.2}$$

respectively. Alice and Bob publicly declare A, B and compute, s_a and s_b:

$$s_a = B^a \bmod p \tag{1.3}$$

$$s_b = A^b \bmod p \tag{1.4}$$

respectively. Since $\Big[(g^a \bmod p)^b \bmod p \Big] = \Big[\big(g^b \bmod p\big)^a \bmod p \Big]$:

$$s_a = s_b = S \tag{1.5}$$

Since at no point, a or b have been disclosed publicly, S constitutes a secret shared key. This secret shared key can then be used in either a symmetric key cryptosystem (described earlier) or as a public key cryptosystem in its own right as follows:

Alice uses A (from above, $A = g^a \bmod p$), g and p as her *public key* information. Given this information Bob can choose a random b, send Alice $g^b \bmod p$ and a message encrypted with $g^{ab} \bmod p$. Since only Alice knows a she is the only person who can decrypt the message.

[1] For every odd prime p, there exists an integer, g, $1 < g < (1-p)$ such that $g^{p-1} \equiv 1 \pmod p$ but $g^n \not\equiv 1 \pmod p$ for every integer n, $1 \leq n \leq (p-1)$. g is a primitive root $\pmod p$

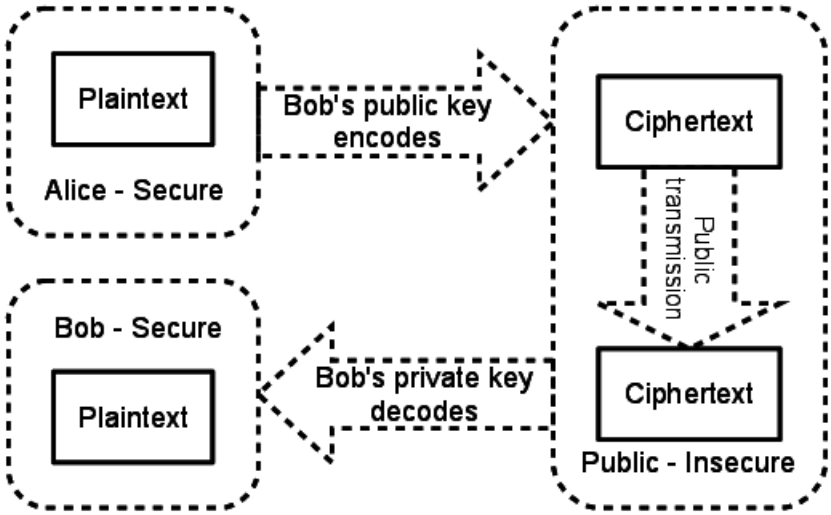


Figure 1.4: Public Key Cryptography Schematic. Alice uses Bob's public key to encrypt a message, the message is decrypted only by Bob's private key.

1.2.2.2 RSA

A more popular public key cryptography protocol is that of Rivest, Shamir and Adleman (RSA) [11]. Generate your public and private key by:

Choose two primes p, q.

$$n = pq \tag{1.6}$$

$$\phi(n) = (p-1)(q-1) \tag{1.7}$$

Choose e, must be coprime[2] to $\phi(n)$ then calculate d where:

$$de \bmod \phi(n) = 1 \tag{1.8}$$

where the public key is n and e and the private key is d. Encryption of the

[2] x and y are coprime if their greatest common divisor is 1. e is often chosen as 65537

plaintext, t, and decryption of the ciphertext, c, are then respectively carried out by:

$$c = t^e \bmod n \tag{1.9}$$

$$t = c^d \bmod n \tag{1.10}$$

The efficacy of this protocol stems from the fact that multiplying two numbers is a trivial operation but finding specific prime factors (without foreknowledge of either factor) is hard. It is, however, unknown as to the exact relationship between the difficulty of cracking RSA and the difficulty of finding prime factors, they are at least equivalent but it is possible that there is an easier way to crack RSA [12]. This is known as the RSA Problem [13]

1.3 Quantum Computing

The two cryptosystems depicted in section 1.2.2 derive their security from problems which are simple to compute in one direction and difficult to compute in the other direction. There are no known *classical* algorithms which, in polynomial time[3], can solve either the integer factorization or discrete logarithm problems.

There is however, a quantum algorithm which, with specific modifications [2], can solve either of the above problems. It uses a classical algorithm to reduce the problem to that of finding the period of a function and then applies the quantum Fourier transform. It is called Shor's algorithm and has been validated experimentally in solid state [14] and more recently on an integrated optical circuit [15]. This is encouraging however current methods have depended on "compiled" systems, i.e.

[3]meaning that the relationship between the algorithm's running time can be expressed as a polynomial dependent on the input size.

experiments specifically designed for a certain factorisation. A universal quantum computer is fortuitously (for those that wish to continue encrypting!) much further away however thought should be invested now as to future prospects and techniques for cryptography.

It is worth mentioning that there is some effort into "post-quantum cryptography" [16] which is a collection of cryptography techniques which rely on problems not solvable by a quantum computer. Little attention will be given to these methods in this thesis since the status of many of these are merely unknown regarding the existence of efficient quantum algorithms. Instead we wish to establish a *guaranteed* solution depending inherently on the laws of physics rather than mathematical computability.

1.4 No Cloning Theorem

Rather than simply running from the ever marching spectre of computational advancement, one might do better to achieve *unconditional security* through some sort of law of nature. This can be done through the use of the "no cloning theorem". [17, 18]

In order to copy a classical bit, one simply needs to measure it and then prepare as many further bits with the same value as the measured bit. This is not the case with quantum bits (qubits) since a quantum measurement is nondeterministic and results in collapse of the wavefunction into a random eigenvalue. It is then necessary to investigate whether one could clone a bit using some operator on a system of a target qubit and a "blank" qubit.

For the sake of argument the blank qubit will be considered to be in the inital state $|0\rangle$. A copying operation, U, of a qubit X onto a blank, e, will then take:

$$U|X\rangle|e\rangle \rightarrow |X\rangle|X\rangle \tag{1.11}$$

$$U|0\rangle|0\rangle \rightarrow |0\rangle|0\rangle \tag{1.12}$$

$$U|1\rangle|0\rangle \rightarrow |1\rangle|1\rangle \tag{1.13}$$

If we try to copy some arbitary state $|\psi\rangle = a|0\rangle + b|1\rangle$:

$$U|\psi\rangle|0\rangle = U(a|0\rangle + b|1\rangle)|0\rangle \tag{1.14}$$

$$= U(a|0\rangle|0\rangle + b|1\rangle|0\rangle) \tag{1.15}$$

$$= aU|0\rangle|0\rangle + bU|1\rangle|0\rangle \tag{1.16}$$

From equations 1.12 and 1.13:

$$aU|0\rangle|0\rangle + bU|1\rangle|0\rangle \rightarrow a|0\rangle|0\rangle + b|1\rangle|1\rangle \tag{1.17}$$

However since:

$$U|\psi\rangle|0\rangle \rightarrow |\psi\rangle|\psi\rangle \tag{1.18}$$

$$|\psi\rangle|\psi\rangle = (a|0\rangle + b|1\rangle)(a|0\rangle + b|1\rangle) \tag{1.19}$$

$$= a^2|0\rangle|0\rangle + ab|0\rangle|1\rangle + ba|1\rangle|0\rangle + b^2|1\rangle|1\rangle \tag{1.20}$$

Comparing equations 1.17 and 1.20, for a and $b \neq 0$:

$$a|0\rangle|0\rangle + b|1\rangle|1\rangle \neq a^2|0\rangle|0\rangle + ab|0\rangle|1\rangle + ba|1\rangle|0\rangle + b^2|1\rangle|1\rangle \tag{1.21}$$

Hence we can copy a specific qubit but not an arbitary one. This raises frustrating issues in quantum computing (such as error corrections) but has some interesting applications to be discussed further in section 1.5.

1.5 Quantum Key Distribution

1.5.1 Seeds of Quantum Cryptography

As discussed in section 1.3, there is clearly a necessity for Cryptography schemes that are resistant to attack by quantum computer, the first steps to this were proposed by Stephen Wiesner in the paper "Conjugate Coding"[4] [20] in which Wiesner first introduces a scheme for transmitting two messages in a manner such that only one may be received. This scheme works on the basis that if the two messages are encoded in the polarizations of photons, message 1 in the horizontal/vertical polarization and message 2 on the left/right circular polarization information. A piece of equipment can only be configured to measure one of the two messages since measuring the circular polarization information destroys any linear polarization information and vice versa so half of the information is lost.

The second scheme Wiesner introduced is for "uncounterfeitable quantum money". in this scheme, a unit of quantum money consists of a collection of two state quantum systems which can be in states a, b, α and β where

$$\alpha = \frac{1}{\sqrt{2}}(a + b) \tag{1.22}$$

[4] a fact not appreciated at the time of its writing, it was written in the '70s and not published until 1983 [19]

and

$$\beta = \frac{1}{\sqrt{2}}(a - b) \tag{1.23}$$

The system then proceeds by encoding each component of the quantum money into one of the four aforementioned states and noting these encodings alongside a conventional serial number. If at any point the unit of quantum money is returned to the quantum mint, the states can be measured and compared to those that were initially encoded.

The uncounterfeitable nature stems here from the fact that a counterfeiter wishes to create an exact copy of the unit of money requires exact information of the states of the quantum systems that make up the unit. His problem lies in the fact that a measurement distinguishing a from b necessarily removes information about whether the state was in α or β. If he attempts to copy the money anyway, there is a 50% chance he attempts to discriminate between the wrong pair of states and thus leaves the money in the wrong state. Assuming the component has been left in the incorrect state, when the mint remeasures the unit[5] there is then a 50% chance that the measurement will be orthogonal to that expected. This leads to a $50\% \times 50\% = 25\%$ chance of discovery. Thus, if the unit of quantum money is made up of a n components, the chance of a successful counterfeiting, p is

$$p = \left(\frac{3}{4}\right)^n \tag{1.24}$$

which, renders a counterfeiting attempt less likely to succeed for increasing values of n. For example, if $n = 20$, $p = 0.00317$ (less than half a percent)

Later, Charlie Bennett and Gilles Brassard [21] showed that encoding informa-

[5]since it prepared the states, it knows which is the correct basis they should be measured in

tion in quantum systems could be used in a general way to share a random secret key between two parties via a Quantum Channel which could then be used for the key in the one time pad (OTP) cryptography scheme. Their scheme relied on using photons as qubits which have the advantage of a very long decoherence time (compared to the time taken to send them between Alice and Bob) which removed the necessity for storing the qubits that was required to realise the concept of Quantum Money.

1.5.2 Protocols

1.5.2.1 BB84

The Bennett Brassard 1984 (BB84) protocol uses the principles of quantum mechanics in order to establish a random secret key known only to two parties, Alice and Bob that is resistant to attacks from an Eavesdropper, Eve [21]. The process is similar to that used in section 1.5.1 where the quantum states, a, b, α and β are implemented as the polarization states horizontal, vertical, diagonal and antidiagonal of single photons, hereby H, V, D, A respectively (figure 1.5).

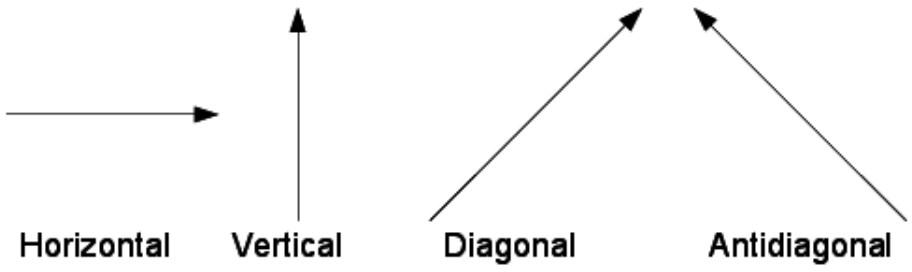


Figure 1.5: The four states used in the BB84 protocol with polarized photons.

These states are related to each other as follows:

$$|\nearrow\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle + |\rightarrow\rangle\right) \tag{1.25}$$

$$|\nwarrow\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle - |\rightarrow\rangle\right) \tag{1.26}$$

$$|\uparrow\rangle = \frac{1}{\sqrt{2}}\left(|\nearrow\rangle + |\nwarrow\rangle\right) \tag{1.27}$$

$$|\rightarrow\rangle = \frac{1}{\sqrt{2}}\left(|\nearrow\rangle - |\nwarrow\rangle\right) \tag{1.28}$$

where the factor of $\frac{1}{\sqrt{2}}$ is necessary to normalise the amplitudes of the states.

Consider the action of a polarizer on these states; the polarizer can be represented by $|\theta\rangle\langle\theta|$, where θ is the polarization orientation.

By this notation we can see that, as $|\langle a|b\rangle|^2$ is the probability b is in state a, the action of a vertical polarizer, hereby denoted $\oplus$ on the horizontal or vertical states are as follows:

$$\oplus|\uparrow\rangle = |\uparrow\rangle\langle\uparrow|\uparrow\rangle = |\uparrow\rangle \tag{1.29}$$

$$\oplus|\rightarrow\rangle = |\uparrow\rangle\langle\rightarrow|\uparrow\rangle = 0 \tag{1.30}$$

Hence we have a test for a photon to analyse whether it is in state $|\uparrow\rangle$ or $|\rightarrow\rangle$ (i.e. horizontally or vertically polarized), if a photon is observed after the polarizer it was V, if it is not observed it was H.

However this is complicated by looking at the action of $\oplus$ on the diagonal/antidiagonal states.

$$\oplus |\nearrow\rangle = \oplus \left(\frac{1}{\sqrt{2}} \Big[|\uparrow\rangle + |\rightarrow\rangle \Big] \right) \tag{1.31}$$

$$= |\uparrow\rangle \langle\uparrow| \left(\frac{1}{\sqrt{2}} \Big[|\uparrow\rangle + |\rightarrow\rangle \Big] \right) \tag{1.32}$$

$$= |\uparrow\rangle \frac{1}{\sqrt{2}} \Big[\langle\uparrow | \uparrow\rangle + \langle\uparrow | \rightarrow\rangle \Big] \tag{1.33}$$

$$= \frac{1}{\sqrt{2}} |\uparrow\rangle \tag{1.34}$$

and the same for $|\nwarrow\rangle$. This means that there is a 50% chance of observing the photon (and 50% chance of it being absorbed) thus there is not a deterministic way to measure an arbitrary polarization. This is a more application specific statement of the *no cloning theorem* discussed in section 1.4.

Using this principle, Alice can send a stream of photons, each randomly encoded in one of the four states from figure 1.5 to Bob who can measure them randomly without concern as to whether he has measured them with a compatible measurement or not. After Bob has signalled to Alice that he has received all of the qubits, Alice then communicates, via a public, but authenticated classical channel, the *bases* each qubit was encoded in. The important thing to note is that she does not reveal the qubit value, only which measurement Bob should have made to know the state with certainty. Alice and Bob can then both discard the values where the transmission and measurement bases do not match leaving them with identical lists without having publicly transmitted them. An example key exchange is displayed in figure 1.6.

The security in this protocol stems from the principle that any measurement on a quantum system will irrevocably modify the system. Consider a modification of the above example wherein an eavesdropper, Eve, intercepts all the qubits, measures

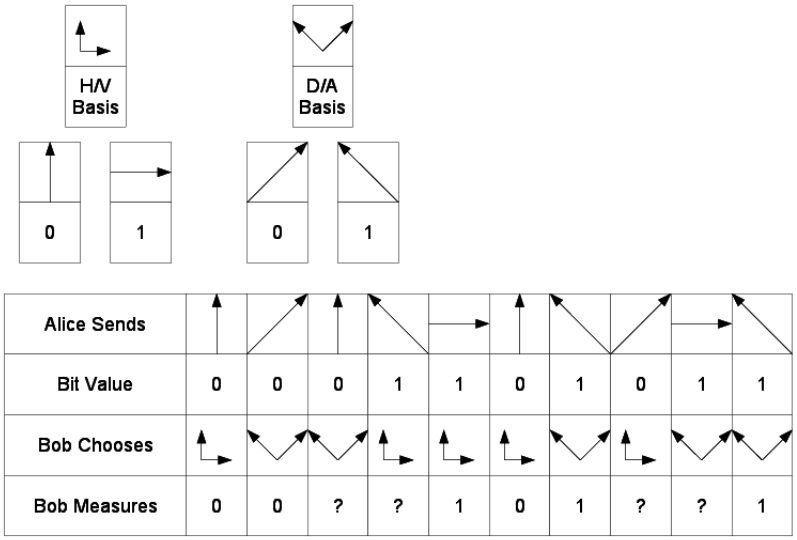


Figure 1.6: An example BB84 key exchange resulting in a secret key of "001011"

them in the same way that Bob would and then, based on her measurements resends these states to Bob. This is called the "intercept-resend attack" and is shown in figure 1.7

Since Bob must have received all of the qubits before Alice reveals what bases she transmitted in, Eve must send the values she measured onto Bob 'blind' and thus with a possibility that she chose the incorrect basis to measure in. There is a 50% chance she will choose the incorrect basis and then, given that she chose the incorrect basis a further 50% chance that the measurement will yield an incorrect result.

Due to this incontrovertible fact of quantum physics, Alice and Bob, upon public discussion, select a subset of their shared string and compare the values of the subset and discard them. The presence of the eavesdropper will be revealed by discrepancies in 25% of the values in their substrings.

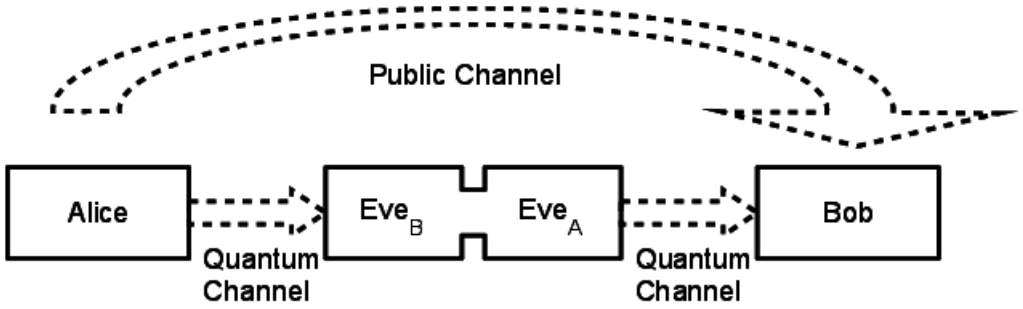


Figure 1.7: Schematic of the Intercept Resend attack. Eve places a detector between Alice and Bob to collect all of Alice's transmissions and then sends the results of her measurements onto Bob. In this case Eve has no control of the public channel.

1.5.2.2 Other Protocols

E91: EPR Based Ekert's E91 protocol [22] is a scheme which uses an entangled system (such as an EPR state [23]) to share qubits between Alice and Bob, keeping pairs where they both made the same measurement on their particle and then tests for eavesdropping by checking for violations of Bell's inequality [24], this protocol has also been demonstrated using photon pairs from parametric downconversion [25, 26] which is equivalent to an EPR system. Some security proofs of BB84 [27, 28] are based on E91 by "replacing all classical randomness by quantum entanglement and postponing all measurements" [29]. This system is useful and some systems [30] have used the principle of two parties measuring an entangled source even though the source was co-located with one of the parties. In spite of the usefulness of this protocol, it requires both parties to have detectors (expensive) and some source of entanglement (expensive and delicate).

B92: Two State Protocol

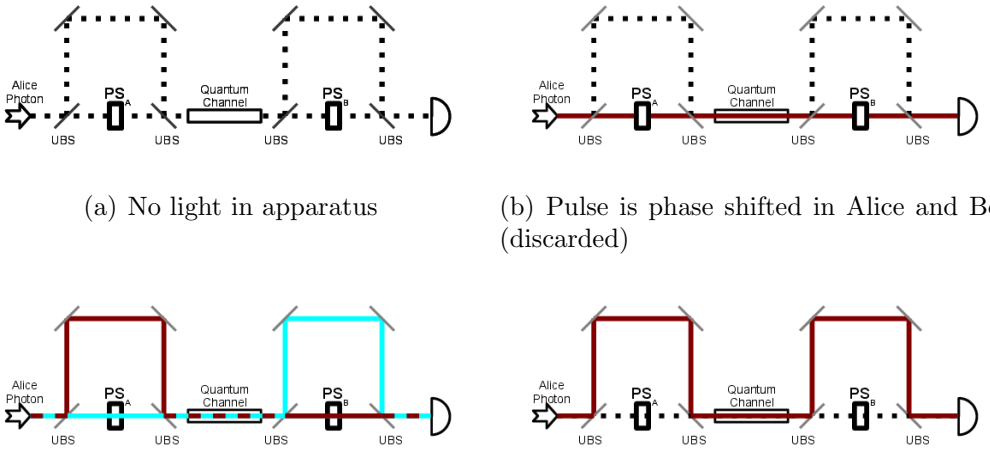
In [31] it is shown that it is sufficient for QKD to only use two nonorthogonal states, encoding one bit value on one state and one on the other. To use the BB84 terminology, this is like encoding $|\uparrow\rangle$ as 1 and $|\nearrow\rangle$ as 0. [31] also suggests

an implementation of the protocol in a fiber based scheme, coding the qubits in phase, which is displayed in figure 1.8. Alice and Bob possess similar apparatus connected via a fiber optic link. The apparatus consist of two asymmetric beam splitters (ABS) - meaning they reflect a far higher proportion of light than they transmit - and a 0/180° phase shifting element in the "dim" (transmission) arm of the arrangement (PS_A and PS_B). When Alice injects a bright pulse into the system, the first beam splitter splits the light into a dim beam to the phase shifter and a bright beam which travels along the long arm of her apparatus picking up a delay of ΔT. These beams are recombined at the second ABS and enter the fiber and into Bob's apparatus. The same arrangement in Bob means that there are three arrival events at Bob's detector.

1: (Figure 1.8(b)). The highly dimmed pulse that travelled through both Alice and Bob's phase shifters arrives. This is discarded as it contains no valuable information.

2: (Figure 1.8(c)). The pulse phase shifted by Alice and delayed by Bob and the pulse delayed by Alice and phase shifted at Bob arrive onto the final UBS simultaneously. Depending on whether Alice and Bob chose the same phase shifts or not, the pulses will interfere either constructively (and be detected) if Alice and Bob chose the same phase shifts or destructively (and not be detected) if they chose different phase shifts. This event generates the key information as Bob can communicate for which pulses a detection was observed without communicating the state of the phase shifter.

3: (Figure 1.8(d)). The twice delayed bright pulse arrives, this contains no key information however does permit Bob to monitor the channel losses relative to the signal pulses and detect eavesdropping.



(a) No light in apparatus

(b) Pulse is phase shifted in Alice and Bob (discarded)

(c) Pulse Phase shifted in Alice delayed by Bob (light line) and vice versa (dark line) (signal)

(d) Pulse delayed by Alice and Bob (monitors eavesdropping)

Figure 1.8: The B92 Protocol performed with phase encoding. Image order as seen from Bob's detections. Fine dotted black lines denote empty channels, light and dark solid lines used to show when two separate light paths are present in apparatus (striped lines denote shared path).

This implementation can be considered as a large interferometer shared between Alice and Bob however since both arms travel in the same fiber, there are no issues with misalignment or unintended asymmetry in the phase changes. This makes it an ideal fiber based system but a lousy free space protocol. This protocol has also been demonstrated in free space [32] where they monitored for Eavesdropping by utilizing two detectors in the receiver and reasoning that an attack which Eve intercepts Alice's signals and forwards very bright pulses when she makes a detection (to counteract the loss that she will introduce by not forwarding all of the signals), the bright pulses will result in a higher than expected rate of coincident counts in the two detectors.

Six State Protocol An interesting refinement over the BB84 protocol is that when a third basis is added [33], this increases the error Eve introduces to the system from 25% to 33% however this is at the cost of decreasing the protocol efficiency from $\frac{1}{2}$ to $\frac{1}{3}$. The level of increased sensitivity to eavesdropping is deemed not sufficiently beneficial in terms of applicability to this work considering both the drop in protocol efficiency and the added complexity of the apparatus that would be required to generate and detect the extra states.

Range extension

Whilst not a protocol *per se*, with imperfect qubit sources and detectors, there exists a limit to the length of a quantum channel. This limit occurs because as the channel gets longer, the losses increase while the noise (dark counts, background) remain the same, thus reducing the signal/noise ratio. At some distance this entirely compromises the transmission as all errors in the sifted key must be considered due to the Eavesdroppper. There exists, however a method of splitting the quantum channel into some number of shorter subchannels over which key can be generated.

This process utilizes the process of *entanglement swapping* [34], wherein Alice and Bob each share an entangled pair of particles with a third party, Charlie. If Charlie performs the correct measurement on his particles, the result will be that Alice and Bob's particles will be entangled. Figure 1.9 shows this (d) and other methods (a-c) for extending the range of the quantum channel.

Unfortunately this process has, thus far, assumed perfect sources and detectors (the source of the problem in the first place!). This is not a problem however since [36] shows that a weak source of entanglement can be "distilled" into a stronger entanglement at the expense of some of the entangled pairs simply by "Local Operations and Classical Communication" (LOCC).

The current limit to introducing quantum repeaters is that since the entangle-

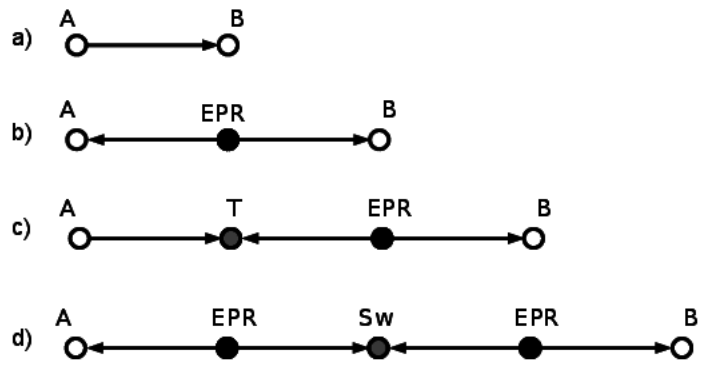


Figure 1.9: Quantum relays with an increasing number of nodes. a) shows a simple link between Alice and Bob. b) shows how the link length can be easily doubled by placing a source of entanglement in the middle of the channel. c) shows a situation where one half of the entangled is sent to "T" who uses it to teleport the state from Alice onto the half of the pair which was sent to Bob. d) shows the situation described in the text where the central point "Sw" posesses one half of Alice's entangled pair and one half of Bob's. "Sw" then performs a measurement to entangle the photons at Alice and Bob. Figure adapted from one in [35]

ment between two nodes may not be established at the same time, it must be stored in a quantum memory until its neighbouring nodes are ready to perform the operations [37]. A candidate for a design is the DLCZ protocol which utilizes atomic ensembles as quantum memories and fairly simple optical components for processing, the efficiency of this protocol scales inversely with distance, an improvement over the non-repeater method which scales exponentially [38].

1.5.3 Attacks

While QKD (and specifically in this case, BB84) theoretically guarantee security, this assumes that all of the components are ideal, this is clearly not the case [39–41]. In order to achieve security, non-ideal components must be identified and analysed to nullify the weaknesses of the implementation.

Intercept-Resend The simplest attack on a QKD system is the *intercept-resend attack*. This is the process mentioned in section 1.5.2 and displayed in figure 1.7 to which QKD is secure against. In brief, Eve intercepts all/part of the qubit stream between Alice and Bob, measures what she intercepted and resends the states she measured. As long as Alice and Bob can communicate with each other, they can detect Eve's presence; they can compare a small, random subset of their results and will observe an increased proportion of bit errors (quantum bit error rate (QBER)) if Eve measured any of the qubits.

Man In The Middle A more powerful attack based on the *intercept-resend attack* is if Eve can convince Alice that she *is* Bob and vice-versa. In this way, two secret keys are generated, k_{AE} between Alice and Eve and k_{EB} between Eve and Bob. In this case, Alice will send her encoded messages to someone she thinks is Bob but actually turns out to be Eve, who can decode the message using k_{AE}, read it and then re-encode using k_{EB}. The man in the middle (MITM) attack is illustrated in figure 1.10.

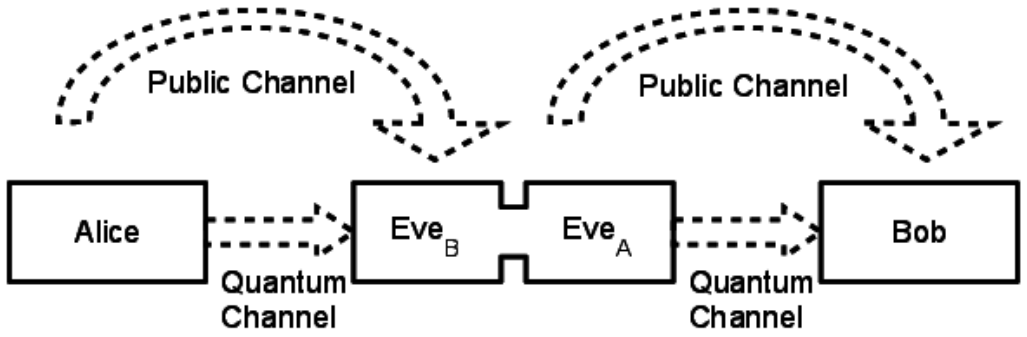


Figure 1.10: Schematic of the man in the middle attack. As in figure 1.7, Eve collects all of Alice's transmission and sends on her measurements to Bob. In this case however she somehow manipulates the public channel to masquerade to Alice that she is Bob and vice versa. In this case Alice and Bob do not generate a shared key with each other, rather they both share separate keys with Eve.

The way in which the MITM attack can be avoided is by use of an *authenticated* channel between Alice and Bob wherein Alice and Bob can be 100% sure they are talking to each other. One way of doing this is to communicate a small amount of a previously agreed secret key to verify their identities. Since this method has shared key as a prerequisite, this requires the idea of quantum key *distribution* to be changed to that of quantum key *expansion*[6].

Side Channels In a recently declassified NSA document from 1972 [42], a range of vulnerabilities were discussed utilizing information leakage from any cryptosystems which could compromise security, this problem was given the codename TEMPEST. Attacks based on this information leakage are called *side channel* attacks. Some examples of exploitable leakage are given below:

- Electromagnetic fields being emitted from a device which could contain key or plaintext information. (initial problem that alerted researchers and caused the initiation of TEMPEST)
- Differing power draw dependent on which parts of the system are active.
- In the case of a system with mechanical parts, the actual sound made by the device could be used to work out certain information. A joke mentioned in [43] is that "the device provided unconditional security against a deaf eavesdropper".

Photon Number Splitting (PNS) It is difficult to manufacture a "true" photon source which probabilistically and deterministically emits one and only one photon on demand [44]. All is not lost however, if a pulsed light source is attenuated such

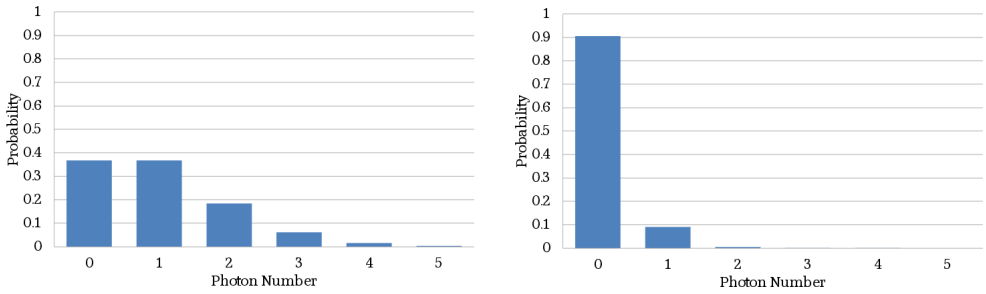
[6]which is actually a problem which also exists in classical cryptography so cannot necessarily be considered as a criticism.

that the average number of photons per pulse is less than one, QKD can still be performed albeit with some added complexity.

The probability an attenuated pulse from a source with a mean photon number μ will contain N photons is given by a Poisson distribution [45].

$$P(N, \mu) = \frac{\mu^N e^{-\mu}}{N!} \tag{1.35}$$

This clearly adds the possibility that the pulse contains more than one photon, which is bad news since it means Eve can split the pulse and measure it without interfering with Bob's measurement. To ensure a low proportion of multiphoton pulses, μ must be much less than one (figure 1.11).



(a) $\mu = 1$ (b) $\mu = 0.1$

Figure 1.11: Probability a pulse from a source with mean photon number μ will contain N photons. Observe that even though single photons are required, $\mu = 1$ produces a large proportion of multiphoton pulses compared to the situation where $\mu = 0.1$

Eve has access to all information about the devices used in the QKD system and this allows her to plan an optimal photon number splitting (PNS) attack. The parameters used in this example are taken from the example in [46] and provide the clearest explanation however in theory any numbers can be used. The attack also depends on Eve possessing a perfectly lossless channel which may sound unrealistic but is not forbidden by the laws of physics so therefore must be considered accessible

to Eve. In this example Alice's pulsed source has a 90% probability of emitting a single photon pulse and a 10% chance of emitting a multiphoton pulse. Now if we assume the channel loss is 90% (such as would be the case in long distance implementations) then Eve's optimal strategy would be to intercept all of Alice's signals, blocking all of those containing only single photons and storing one photon from each multiphoton pulse while forwarding the other(s) along a lossless channel to Bob. From Bob's point of view, 10% of the signals are still reaching him so suspicion is not aroused however all photons received have a copy possessed by Eve who can retain them[7] and measure them once Alice and Bob have publicly discussed their bases.

This attack places a constraint on the maximum photon number of the source:

$$y > p_m \tag{1.36}$$

$$y > (1 + \mu)\, e^{-\mu} \tag{1.37}$$

where y is the *yield* of the quantum channel (proportion of the signals which are detected) and p_m is the probability that the source emits a multiphoton pulse. This means that as the yield decreases (loss increases), less multiphoton pulses can be tolerated.

A countermeasure to this attack which removes the dependence on channel loss is discussed later section 1.5.3.1. Due to the immaturity of true single photon sources, the QKD security proofs (section 1.6) consider the use of imperfect devices such as attenuated pulsed sources and the value of μ can be optimised (figure 1.12) to provide the highest secret key rate.

[7] again, Eve must be afforded any operation which is not forbidden by the laws of Physics

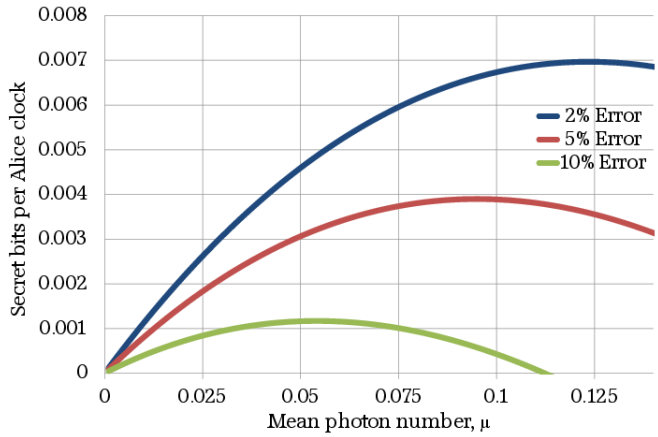


Figure 1.12: A graph of the rate (per transmitted bit) of BB84 for various photon numbers indicating there is an optimum value of μ dependent on the system parameters. In this example the yield is 10%. This is calculated from the GLLP proof discussed in section 1.6.

Indistinguishability If different devices are used for generating each qubit state, then great thought must be given to the indistinguishability of the photons, if, for example, there is a large difference between the wavelengths of each source, then a dichroic beamsplitter could be used to determine, with precision, which source, and consequently which qubit state the photon was sent in. This must be dealt with by control of spectral intensities of the sources and tight spectral filtering at the output of the Alice optical device.

If the weakly pulsed source is a laser, then there is phase information in the pulses which could be exploited for an attack. In order for this to be counteracted, active phase randomization must be employed. A simpler solution is to use an incoherent source such as an LED.

Back Flash An issue based on detection is due to the fact that the carriers in a single photon avalanche diode (SPAD) can recombine [47] a short time after an avalanche creating a *back flash* which, in the case of the 4 detector BB84 system would give information as to which state Bob measured [39]. Mercifully, this re-emitted light intensity is not evenly spread across the spectrum and so the signal

photon wavelength can be chosen such that narrow spectral filtering can be employed to only allow the transmission of signal photons and blocking the back flash photons (from [39], filtering all wavelengths above 700nm should suffice).

Blinding Another detector based attack is one which utilizes the imperfection of the detectors themselves. Makarov [48, 49] discovered that, using a laser beam of specific power, a SPAD can be put into "saturation" mode until the laser beam intensity is reduced, de-saturating the SPAD whereupon it will produce a signal as if it had just detected a photon. Using several laser beams at various polarizations and intensities, it is possible to saturate all detectors in a passive BB84 system and cause single detectors to click at will. This attack makes Eve far more powerful when performing the intercept-resend attack since she can make it such that Bob will detect the same state that she did with certainty by choosing the appropriate laser pulses. When Alice then transmits the bases, Eve and Bob will end up keeping the same qubit measurements and thus will end up with the same key.

The generality of this attack is still under debate; [50] suggests various countermeasures and situations where the attack may not work.

This attack has also been experimentally carried out against an established system [51]. The authors also recommend further countermeasures which could neutralise the attack [52].

1.5.3.1 Improvements on BB84

Further enhancements can be made to the basic BB84 protocol discussed in section 1.5.2.1 for enhanced security or key rate. The modifications discussed below consist of, at worst, trivial modifications to the practical setup of a BB84 system.

Basis Choice Bias Due to the random basis choice performed at Alice and Bob, the inherent efficiency of the protocol discussed above is at best 50%. Ardehali et al. [53] suggested that if Alice and Bob use a pre-chosen bias to one basis or the other, their measurements will agree far more than if they were to choose a simple 50:50 basis choice. This scheme, however, opens an obvious issue that since in the security proof of QKD, one has to assume that Eve knows everything about Alice and Bob's devices, if this were the case then Eve could simply eavesdrop only the basis with a high choice bias and gain a large portion of information of the key without disrupting many of the qubits due to her high probability of making the correct bias choice.

The solution to this issue is simple, Alice and Bob must, instead of comparing the QBER for the entire transmission, compare the QBER separately for each basis.

Consider the case when Alice and Bob choose the two bases with probability p and $1-p$, where $0.5 < p < 1$. Eve's optimal strategy would be to eavesdrop only in the p basis. Under normal error checking, this would introduce an error of $\frac{(p-1)}{2}$, which, in the case that p becomes asymptotically close to 1 would result in a negligible error contribution.

If, however, the QBER is calculated separately for each basis, the QBER in the p basis will be 0 since Eve is always choosing the correct basis however since Eve's optimal strategy for eavesdropping is *always* measuring in the incorrect basis in the $1-p$ channel, there will be a 50% QBER in the other channel.

Obviously, as p moves away from 0.5 the number of bits available for error estimation decreases and thus Eve has a better chance of eavesdropping undetected, a suitable value for p must be chosen. It has been shown [54] that the optimal p scales with key length by:

$$p = \mathcal{O}\left(\sqrt{\frac{\log k}{N}}\right) \tag{1.38}$$

where k = length of final key, N = number of qubits shared. Meaning that as $N \to \infty$, $p \to 0^+$.

This result is not taken into consideration in the GLLP security proof (section 1.6.1) since it deals with the asymptotic limit but it is taken into consideration in the finite-key proof (section 1.6.2) which deals in general with the issue of having sufficient bits to estimate the QBER accurately.

Decoy States [46] introduces the concept of *decoy* pulses, that is, introducing a number of pulses from a weak coherent pulse (WCP) source that have a high μ so that it emits a higher proportion of multiphoton pulses. This protocol significantly improves BB84 protocol against the PNS attack discussed in section 1.5.3.

The Decoy State protocol introduces two sets of states to be emitted by Alice, those with mean photon number μ and those with μ'. Now, after the transmission, Alice can publicly reveal, along with the basis as in standard BB84, whether the pulse was from the decoy source or the signal source. Since Eve's optimal intercept-resend strategy discussed in section 1.5.3 is to block all single photon pulses and split the multiphoton pulses, the losses in each case can then be compared and if the losses of signal pulses are higher than the losses in decoy pulses then eavesdropping is detected and Eve is foiled.

The condition for security from [46] is thus:

$$\frac{e^{\mu'}}{\mu'}\frac{\mu}{e^{\mu}} < 1 \tag{1.39}$$

A striking feature of equation 1.39 is that this criterion does not depend on the channel loss, meaning security is maintained for a given system regardless of the loss (eg, changing channel distance).

Detailed security analysis of this protocol can be found in [55, 56].

SARG04 In a 2004 paper, Scarani et al. [57] proposed a modification to the BB84 Protocol, differing only in the classical sifting stage allowing this new protocol to be trivially implemented on an existing BB84 system. The protocol will hereby be referred to as SARG04 and a brief description will follow.

For simplicity in this specific explanation, a slightly different notation will be employed. Here, the four states transmitted by Alice will be denoted by $|\pm x\rangle$ and $|\pm z\rangle$ and the measurements made by Bob denoted by σ_x and σ_z where:

$$\sigma_x |+x\rangle = +1 \tag{1.40}$$

$$\sigma_x |-x\rangle = -1 \tag{1.41}$$

$$\sigma_x |+z\rangle = \pm 1 \tag{1.42}$$

$$\sigma_x |-z\rangle = \pm 1 \tag{1.43}$$

and the inverse for σ_z.

As in "standard" BB84, Alice chooses randomly one of the four states from $|\pm x\rangle, |\pm z\rangle$ and Bob chooses randomly one of the measurements σ_x, σ_z. Now, instead of communicating the basis Alice sent in, ie transmitting $|\pm x\rangle$ or $|\pm z\rangle$ with one state in each basis corresponding to 0 or 1; Alice announces to a pair of nonorthogonal

states to which the state belongs, the pairs being:

$$\mathcal{A} = \begin{cases} |+x\rangle, |+z\rangle \\ |+x\rangle, |-z\rangle \\ |-x\rangle, |+z\rangle \\ |-x\rangle, |-z\rangle \end{cases} \tag{1.44}$$

henceforth referred to by $\mathcal{A}_{w,w'} = \{|wx\rangle, |w'z\rangle\}$.

The second difference is the correspondence between states and key bit values with $|\pm x\rangle$ corresponding to 0 and $|\pm z\rangle$ corresponding to 1. Bob can then compare his obtained measurement with the possible results his chosen measurement will produce if enacted on $\mathcal{A}$ to see whether his obtained measurement allows definite knowledge of the received state.

Figure 1.13 shows a sample bit transmission and the range of measurement outcomes. If Alice sends $|+x\rangle$ and Bob measures σ_x he will definitely measure +1. However, since $\sigma_x |\pm z\rangle$ can also yield +1, whichever $\mathcal{A}$ Alice announces Bob will have to throw away his result. (rows a and c)

If Bob measures σ_z however, he will measure either +1 or -1 with 50% probability each. If Alice announces $\mathcal{A}_{++}$ then since $\sigma_z |+z\rangle = +1$, if Bob measured -1 then he knows that the state he measured must have been $|+x\rangle$ (row b). Conversely, if Alice announces $\mathcal{A}_{+-}$ then Bob knows with certainty that a +1 result from a σ_z measurement corresponds to the $|+x\rangle$ state (row c).

If one completes the analysis for the instances when Alice sends $|-x\rangle$ and $|\pm z\rangle$ then there is an extra factor of $\frac{1}{2}$ due to the fact that even if Bob choses the correct basis (which, as in BB84 he will do half of the time) there is still a further $\frac{1}{2}$ chance that the measurement will not yield a usable bit. This extra factor of $\frac{1}{2}$ reduces the protocol efficiency for a given photon number down to $\frac{1}{4}$

Alice Sends	Alice Declares	Bob Measures	Bob's Result	Possible Results on $\mathcal{A}_{w,w'}$	
|+x>	$\mathcal{A}_{++}$	σ_x	+1	$\sigma_x\|+x\rangle = +1$	a)
				$\sigma_x\|+z\rangle = +1$ or -1	
		σ_z	+1 or -1	$\sigma_z\|+x\rangle = +1$ or -1	b)
				$\sigma_z\|+z\rangle = +1$	
	$\mathcal{A}_{+-}$	σ_x	+1	$\sigma_x\|+x\rangle = +1$	c)
				$\sigma_x\|-z\rangle = +1$ or -1	
		σ_z	+1 or -1	$\sigma_z\|+x\rangle = +1$ or -1	d)
				$\sigma_z\|-z\rangle = -1$	

Figure 1.13: An example transfer of a single bit using the SARG04 sifting method. In situation *a)*, Bob measures σ_x on $|+x\rangle$ and measures +1 with certainty, Alice declared that she either sent $|+x\rangle$ or $|+z\rangle$. Since σ_x of either of these can give a measurement outcome of +1, Bob cannot be sure whether he detected$|+x\rangle$ or $|+z\rangle$. In situation *b)* on the other hand, if Bob measures σ_z, he will measure +1 or -1 with equal probability, in the case that he measured +1, as in *a)* either of $|+x\rangle$ or $|+z\rangle$ could have caused that outcome however the only way -1 could have been measured is if the state was $|+x\rangle$ (which occurs half of the time σ_z is performed on $|+x\rangle$. The same logic can be applied to lines c and d of the figure and for $|+x\rangle$ and $|\pm z\rangle$ (not shown) to compute all of the possible permutations.

It is shown in further detail in [58] that the SARG04 protocol is secure against PNS attacks under conditions where BB84 is provably insecure for both single photon sources and WCP sources. It also compares the performance of SARG04 and BB84 for a range of distances wherein BB84 is found to be better at short distances (ie when losses are still small so the protocol efficiency is the dominant factor) and therefore whilst it could be of academic interest, SARG04 is not suitable for obtaining QKD over the short ranges in this thesis.

1.6 Security Proof

The security of QKD is guaranteed only if idealized devices are used so any real world application needs to be analysed based on how much information it leaks to a theoretical eavesdropper. To discuss this a basic QKD process will be laid out.

Alice prepares data The first step for the transfer is for Alice to generate a random string of bit pairs which can be converted into polarizations states of qubits by, eg. $00 \Rightarrow |\uparrow\rangle$, $01 \Rightarrow |\rightarrow\rangle$, $10 \Rightarrow |\nearrow\rangle$ and $11 \Rightarrow |\nwarrow\rangle$. As long as Alice and Bob agree on this mapping it can be any permutation of this.

Transmission Alice then transmits weak pulses with the polarization states she generated. Bob randomly chooses the H/V or D/A basis to measure each photon and records the times of the detections.

Raw Key Generation Since every pulse from Alice did not contain a photon and Bob's detectors are not 100% efficient. Bob now possesses a sparse list of detections and their measurement outcome. Bob then shares some subset of his sparse list with Alice which is used to estimate the QBER and determine the synchronisation offset between the data sets.

Error Correction Even in the absence of eavesdropping, due to imperfect state preparation and measurement and also dark counts in the detectors, there will be a nonzero number of errors in the sifted keys which need to be removed for the keys to be any use. An error correction code is utilized to recursively check for errors without revealing the key information.

Privacy Amplification Some sort of hashing is performed to reduce the amount of information a theoretical eavesdropper could have on the key to zero. Although in the previous point it is suggested that there are errors introduced from a

non-eavesdropping source, these errors must be attributed to the eavesdropper anyway and therefore constitute information leaked.

This section will concern itself with the amount of information leaked and, combining that with the mean photon number and the channel losses, estimate how many secret bits per qubit sent by Alice will be generated.

1.6.1 Infinite bound

As discussed in section 1.5.2 there is a threshold QBER of 25% at which it must be assumed that the quantum transmission has been compromised and any key generated is not secure. In real world applications of QKD however, this is not Eve's optimal strategy for eavesdropping as she can utilize the imperfections in the devices in the system to gain information on the key without increasing the QBER. In this case the no cloning theorem discussed in section 1.4 is no longer sufficient and a more rigorous security analysis must be undertaken. The result from this analysis must then be used to optimize the parameters of the apparatus and discern the maximum tolerable QBER.

The quantum channel is "noisy" Any errors resulting from noise/loss in the quantum channel must be attributed to Eve since she could, in theory, replace the channel with a quieter/less lossy one and reintroduce the noise/loss by way of eavesdropping.

Alice does not use a single photon source It must be assumed that all multi-photon pulses emitted by a WCP Alice are intercepted by Eve and used to gain key information.

Bob's detectors are not perfectly efficient If Eve had the technology, she could modify Bob's detectors to be more efficient and then attack sufficient photons such that Bob sees no change in detections.

Expanding work carried out by [27, 59–61] assuming various imperfect device scenarios, the GLLP proof [62] introduced a second attacker, Fred, who has the power of tampering with the source and detector.

GLLP also formulated an expression for the maximum key generation rate given a certain set of component parameters

$$R = max\left((1-\Delta) - H_2(\delta) - (1-\Delta)\,H_2\left(\frac{\delta}{1-\Delta}\right)\right) \tag{1.45}$$

where Δ is the probability that a qubit is 'tagged'[8], δ is the QBER and H_2 is the Shannon Entropy function [7] defined by:

$$H_2(x) = -x\log_2(x) - (1-x)\log_2(1-x) \tag{1.46}$$

equation 1.45 can be rewritten in terms of easily measurable parameters of the experiment [63] as:

$$R \geq q\left(-Q_\mu H_2(E_\mu) + Q_1\left(1 - H_2(e_1)\right)\right) \tag{1.47}$$

where

$$q = \frac{N_s^\mu}{N} \tag{1.48}$$

[8] a qubit that has leaked information to Eve

where N_s^μ is the number of signals for which Alice and Bob used the same basis and N is the total number of pulses; in BB84, for sufficiently large N, $q = \frac{1}{2}$.

$$Q_\mu = \frac{K_s^\mu}{N_s^\mu} \tag{1.49}$$

where K_s^μ is the number of detections made by Bob, making Q_μ the probability a given occasion upon which Alice and Bob agree on the basis will result in a bit of sifted key, this can be calculated by considering the mean photon number of the source and the efficiency of Bob's detectors;

$$E_\mu = \frac{K_s^{error}}{K_s^\mu} \tag{1.50}$$

where K_s^{error} is an incorrect bit in the sifted key, making E_μ the QBER, which is directly measurable from the system by sending known states and measuring Bob's output;

$$Q_1 = Y_1 \mu e^{-\mu} \tag{1.51}$$

where Y_1 is the chance Bob will make a single photon detection given Alice emits one and μ is the mean photon number of the WCP source, or, as per [63] for non-decoy BB84:

$$Q_1 = Q_\mu - p_M \tag{1.52}$$

where p_M is the probability of Alice sending a multiphoton state (see equations 1.36 and 1.37). Finally, the error rate of a single photon state e_1 [63]:

$$e_1 = \frac{Q_\mu E_\mu}{Q_1} \tag{1.53}$$

Substituting these specific expressions for Q_1 and e_1 into equation 1.47

$$R \geq q\left(-Q_\mu H_2\left(E_\mu\right) + \left(Q_\mu - p_M\right)\left(1 - H_2\left(\frac{Q_\mu E_\mu}{Q_\mu - p_M}\right)\right)\right) \tag{1.54}$$

In addition to giving an expression for the optimal mean photon number μ as in figure 1.12. It also shows an interesting result that the optimal μ will never exceed η and will only tend towards η as the QBER tends towards zero (shown in figure 1.14 as an unrealistic QBER of 0.1%[9]). This is because of the information leakage in the PNS attack.

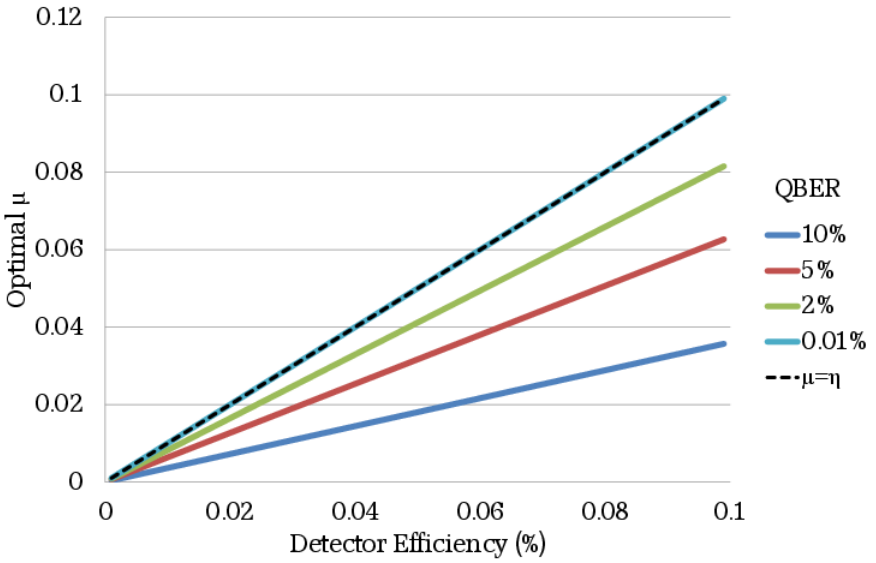


Figure 1.14: The optimal mean photon number, μ, depends on the yield, η, and the QBER however $\mu < \eta$.

[9] $H_2(0)$ is undefined

1.6.2 Finite Key

The problem with the GLLP proof (and those which preceded it) is that the QBER is required to determine the amount of privacy amplification necessary for an unconditionally secure key. This QBER is estimated by using a small subset of the sifted key, which, tends towards the real QBER for an infinite key however for a key of finite length this QBER needs to be upper bounded and, since Eve can get information "for free" on any multiphoton pulses, the single photon yield (proportion of pulses which contained a single photon) needs to be lower bounded also.

Using the formulation of conditional entropies of Devetak and Winter [64] which expresses the rate as:

$$r = H(X|E) - H(X|Y) \tag{1.55}$$

where $H(a|b)$ are conditional von Neumann entropies (which become Shannon entropies when the systems are classical) signifying the mutual information possessed by (Alice and Eve) and (Alice and Bob), respectively.

For the sake of completeness, for infinite key, this becomes [65]:

$$r = p_z^2 \left(S(X|E) - leak_{EC} \right) \tag{1.56}$$

where:

$$S(X|E) = 1 - Y_1 \left(1 - h\left(\frac{Q}{Y_1} \right) \right) \tag{1.57}$$

$$Y_1 = 1 - \left(\frac{\tilde{\nu}_S}{R} \right) p_m \tag{1.58}$$

where R is the total detection rate, $leak_{EC}$ is the amount of information leaked

during error correction, Q is the total QBER, $\tilde{\nu}_S$ is the transmission rate of Alice divided by the proportion of detections make up the sifted key and p_m is the proportion of Alice's pulses which contain more than one photon.

The problem discussed in [66] is that $S(X|E)$ needs to be expressed in terms of a security parameter which expresses the precision to which it is estimated. The result [29] is that:

$$r = \frac{n}{N}\left(S_\xi(X|E) - \Delta - \frac{leak_{EC}}{n}\right) \tag{1.59}$$

where n is the fraction of the data N which is used for key, $S_\xi(X|E)$ is the estimate of Eve's information, Δ is the security parameter associated with privacy amplification:

$$\Delta = 7\sqrt{\frac{log_2(\frac{2}{\bar{\epsilon}})}{n}} + \frac{2}{n}log_2\left(\frac{1}{\epsilon_{PA}}\right) \tag{1.60}$$

$$leak_{EC} = 1.05h(e) \tag{1.61}$$

where $\bar{\epsilon}$ is the security parameter associated with the parameter estimation and ϵ_{PA} is the security parameter associated with privacy amplification.

For BB84 without Decoy States, $S_\xi(X|E)$ is given to be [67]

$$S_\xi(X|E) = \tilde{Y}_1(\mu)\left(1 - h(\tilde{e_x}(1))\right) \tag{1.62}$$

where

$$\tilde{Y_1}(\mu) = \frac{1 - p_m}{\mathbf{R}} \tag{1.63}$$

$$\tilde{e_x} = \frac{e_x}{\tilde{Y_1}(\mu)} \tag{1.64}$$

A striking difference between the finite key method and the asymptotic method is that the finite key method predicts negative key rates for $N \lesssim 10^5$ [66] and only arrive at the asymptotic values for $N \approx 10^{10}$.

1.7 Free Space QKD

1.7.1 Experimental History

The first experimental realisation of QKD was presented in a paper by Bennett, Bessette, Brassard, Salvail and Smolin in 1992 [19] and consisted of a pulsed green LED which was collimated, filtered and horizontally polarized before being rotated by a pockels cell which can rotate the horizontally polarized light into any of the four states required in the BB84 protocol (this experiment used H, V, L and R polarizations). Following a 32cm free space channel a second pockels cell set to perform no rotation or to rotate L/R to H/V (and vice versa), a polarizing beam splitter discriminating between H/V (therefore L/R if the second pockels cell was active) directed the light onto photomultiplier tubes. A photograph of this apparatus can be seen in figure 1.15. A sample data collection run is provided in the paper which details that 715,000 pulses were sent by Alice in 10 minutes, an average repetition rate of 1.1kHz.

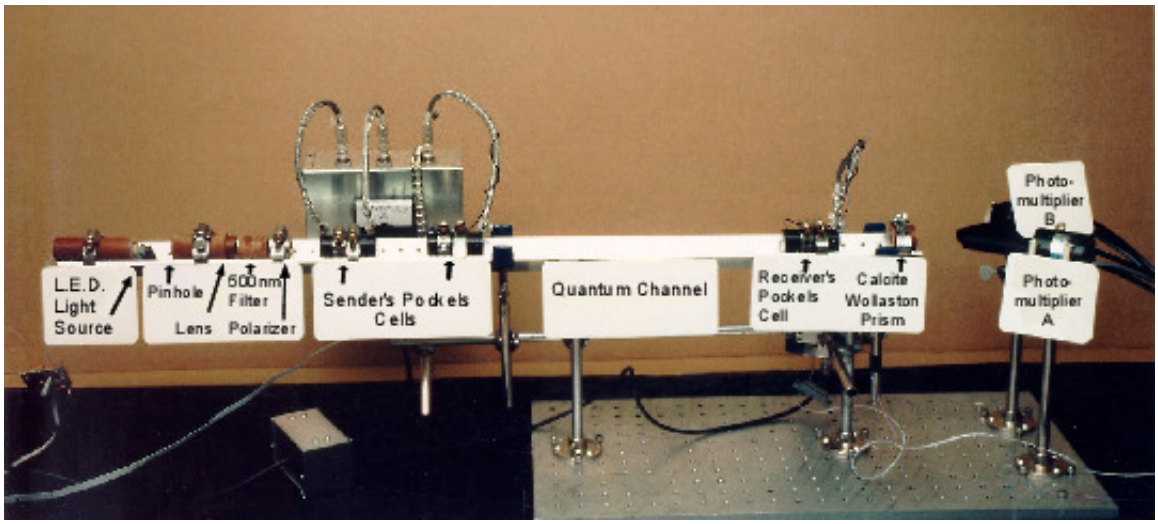
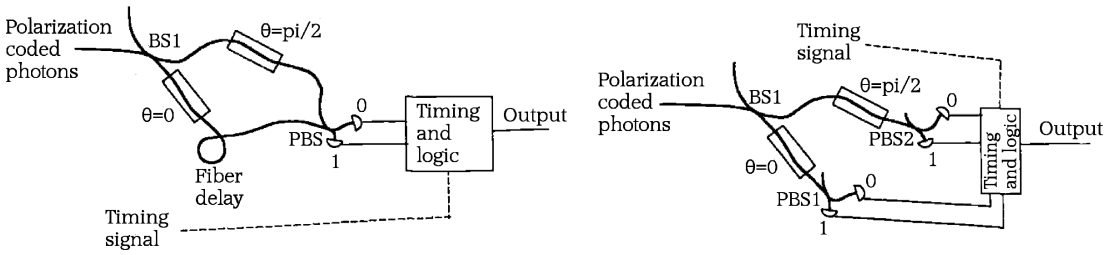


Figure 1.15: A photograph of the apparatus used in the BBBSS demonstration of QKD. Image taken from [19], original labelling edited for clarity.

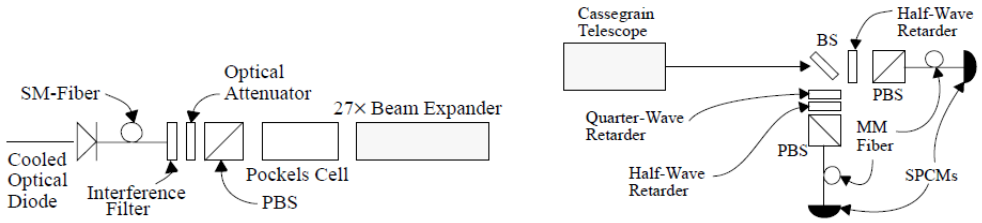


(a) Two detector passive QKD receiver (b) Four detector passive QKD receiver

Figure 1.16: The methods of randomly choosing measurement bases proposed by Rarity, Owens and Tapster. Images taken from [68]

The methods of randomly choosing transmission and measurement states were refined in a paper by Rarity, Owens and Tapster in 1994 [68] which proposed replacing the pockels cells with beam splitters and either placing a polarizing beam splitter (PBS) on the transmitted and reflected arms (figure 1.16(a)) or putting a fiber delay on one arm and recombining after a delay (figure 1.16(b)) retaining the ability for two detectors to discriminate between the four states. The BBBSS experiment required $60\mu S$ to stabilize after each rotation whereas this paper rightfully points out that "As the beam splitter acts passively, there is no maximum switching rate and the data rate is limited only by the detector time resolution"[10]

In a 1998 paper Buttler et al. [32] at Los Alamos demonstrated QKD over 1km of free space at night time. Their system maintained a pockels cell for state prepa-



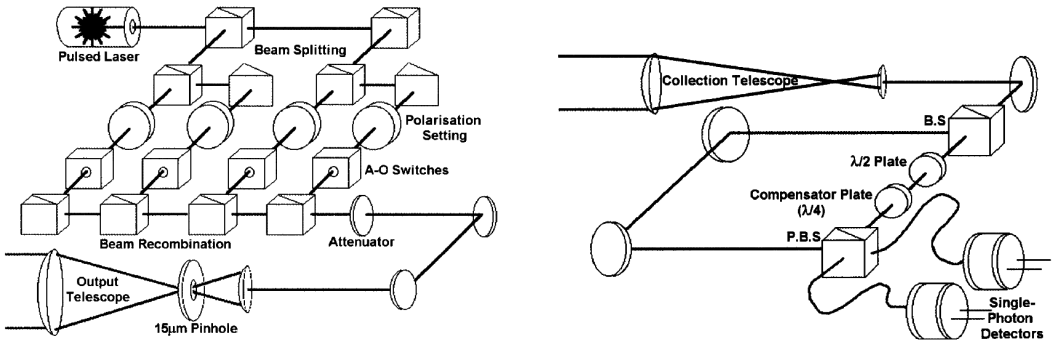
(a) Transmitter (b) Receiver

Figure 1.17: The apparatus used in the 1km Los Alamos experiment. Images taken from [32]

[10] A direct quote from section 4 of the paper

ration (figure 1.17(a)) but used a 50:50 beam splitter in the detection process (figure 1.17(b); the Bennett 1992 (B92) protocol was utilized but the principle is the same). The transmission rate detailed in this paper is 20kHz.

Rather than rotating the polarization of a single beam, Rarity et al. [69] exchanged key over 1.9km by using a set of beam splitters to generate four beams from one laser, polarized each individually and then used acousto-optical switches to select which of the four beams would be routed to the output of the transmitter (figure 1.18(a). The detection system used the time multiplexed version of the beam splitter design from [68](figure 1.18(b)).



(a) Transmitter (b) Receiver

Figure 1.18: The apparatus used in 1.9km experiment of Rarity et al. Images taken from [69]

A further result from Los Alamos by Hughes et al. in 2002 [70] extended the distance to 10km in day and night conditions. This system utilized four polarized diode lasers being combined using 50:50 beam splitters and symmetrically the detector comprised a 50:50 beam splitter and two PBS as described earlier (figure 1.19). The clock rate in this system was also increased to 1MHz.

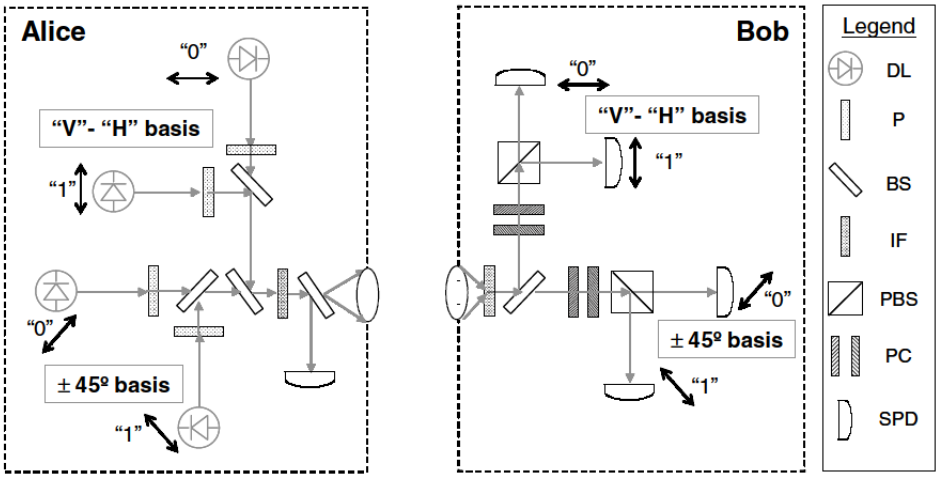


Figure 1.19: The apparatus used in the 10km Los Alamos experiment. Images taken from [70]

Also in 2002, Kurtsiefer et al. [71] performed a 23.4km experiment at night between two mountains in Germany. This system used the beamsplitter design in its detection module however implemented a novel design in combining the polarized laser diode beams. The beams were arranged around a conical mirror and spatially filtered (figure 1.20). This paper also reported performance in adverse conditions and reasoned that the performance meant that ground to satellite transmissions were feasible.

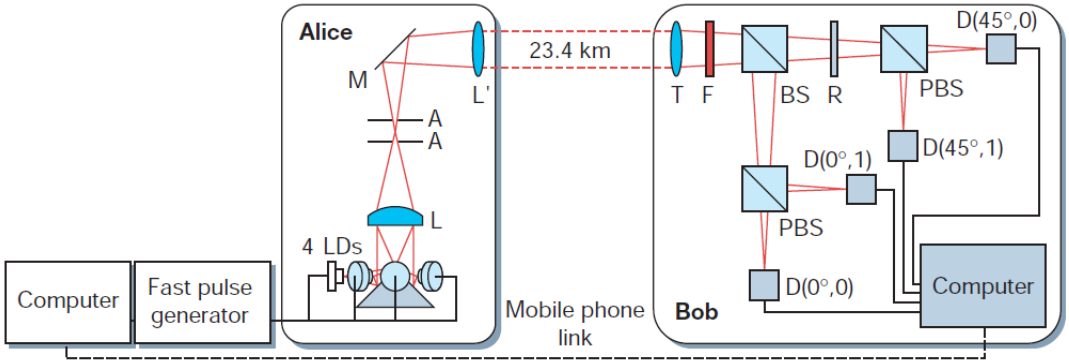


Figure 1.20: The apparatus used by Kurtsiefer et al. for the 23.4km experiment. Image taken from [71]

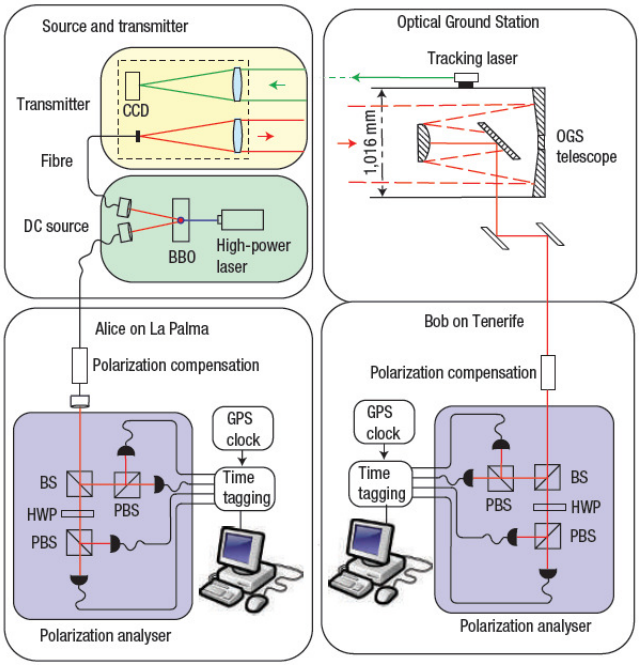


Figure 1.21: The apparatus used by Ursin et al. for the 144km experiment. Image taken from [30]

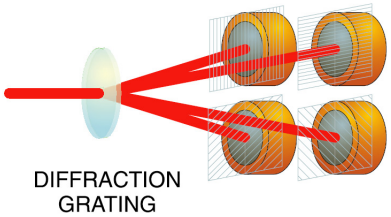
The current furthest free space QKD link (144km) was reported by Ursin et al. in 2007 [30] between La Palma and Tenerife. This process actually used an entanglement protocol wherein both parties must receive qubits (see figure 1.21). Shor's security proof [27] proves that this is equivalent to BB84 and even so, much of the required apparatus is the same anyway. The beam wander from the f/38 telescope was much larger than the size of the detectors so an f/5 beam reducer was employed to reduce the wander down to less than $500\mu m$. To this end a detector arrangement was devised which placed a 50mm lens before the first beam splitter however this was not completed in time and a scheme utilizing 50mm lenses in front of each detector was implemented instead.

The unused long range detector unit has evolved into the f/5 design utilized in this thesis. The f/5 optics is not strictly necessary (the beam wander over 10cm is negligible) however it does mean that it would be relatively simple to re-purpose the "Quantum ATM" unit for long range QKD experiments.

1.7.2 The Bristol System

What follows is a brief history of the Bristol low cost QKD system and its evolution to the current specifications. Before the f/5 design was integrated into the Bristol short range system, a compact detection system was employed which was shown by Duligall et al. [72]. In this system, four LEDs were combined onto a two dimensional diffraction grating (hereby a diffractive optical element (DOE)) and then spatially filtered with pinholes. In a symmetric arrangement, the collimated beam sent across the quantum channel was split by a further DOE onto one of four detectors each with a polarizer across them such that each detected one of the four BB84 polarizations. This reduces the protocol efficiency by $\frac{1}{2}$ for the sake of size and simplicity but in the end it was decided that since the detection unit will reside in the "Quantum ATM", size is not really of issue and, in fact, due to available modifications [54] to the BB84 protocol (discussed in section 1.5.3.1) it is desirable to be able to choose the weighting of the random basis choice which requires the beam splitter arrangement.

(a) DOE beam combiner



(b) DOE beam separator

Figure 1.22: DOE beam combiner and separator used in the first Bristol system. Images from [73]

1.7.3 The system as of the commencing of this thesis

The work detailed in this thesis commenced October 2007. At this point the system consisted of Alice and Bob optical devices each bolted to an optical bench figure 1.23. The Alice optics were as described in section 1.7.2, four LEDs incident on a diffraction grating to collimate and filtered by pinholes. The Bob optics inherited was a design of the sort shown in figure 1.19 and figure 1.20, namely using a 50:50 beamsplitter for basis choice and a PBS for each measurement in the bases. This design placed each PBS on a tip/tilt platform however this meant that the beams transmitted by each PBS needed to be aligned simultaneously.

To analyse the timing, a GuideTech GT653 time interval analyser (TIA) was utilized, this only had two channels so in each basis one signal was delayed by half a clock cycle and combined with the undelayed pulse. This had the effect of doubling the apparent background in each channel.

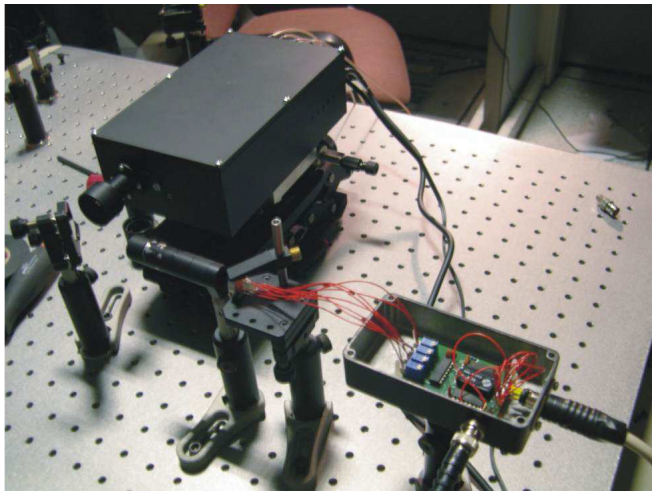


Figure 1.23: A photograph of the Alice (bottom right) and Bob (top left) devices taken from [73]. Interfacing and processing was performed on a desktop PC (not shown).

1.8 Commercial Systems

Quantum Cryptography is arguably the first commercial application of quantum information research and there are currently several companies offering QKD solutions [74–76]. An example system is the Clavis 2 produced by idQuantique depicted in figure 1.24. While these are perfectly appropriate for applications wherein the large size and considerable cost are of little issue (government, military), there are currently no small or low cost systems suitable for the consumer market. One might consider this to be of little consequence if one compares the relative sensitivity of the data an average person might want to communicate to that of a country or large multinational, and to a certain extent that is true for the time being however section 1.2 has elucidated the necessity for a widespread implementation for QKD in most situations where classical cryptosystems are currently used.

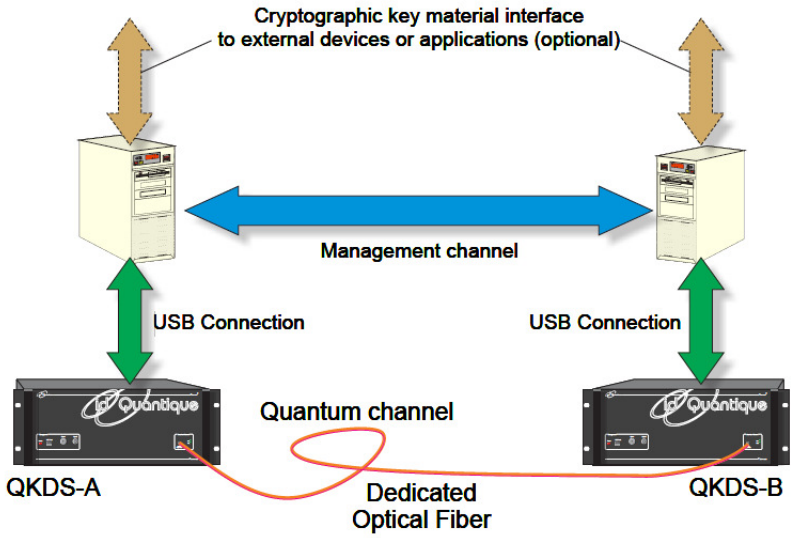


Figure 1.24: Clavis 2 QKD system from IDQuantique (image taken from [74])

1.9 Use Scenarios

The general idea for a consumer QKD system is depicted in figure 1.25 wherein several small devices can all establish key with a centralised terminal which can then be used for some kind of cryptographic application. Two possible applications are described here.

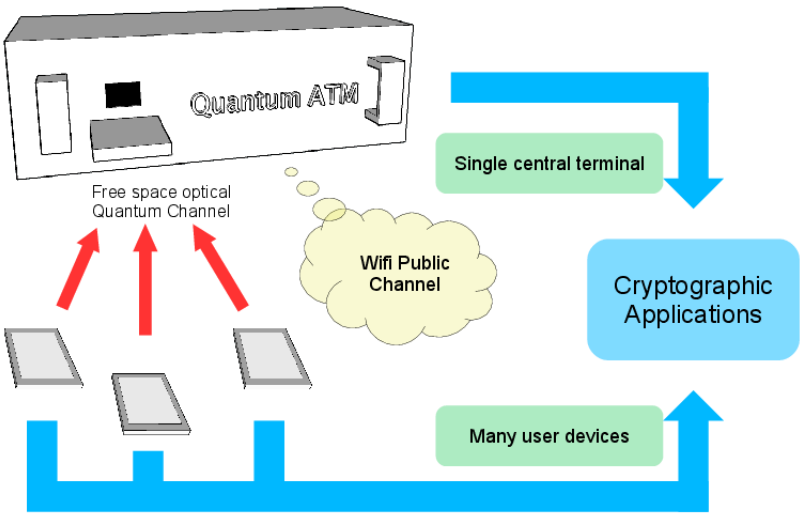


Figure 1.25: Cartoon of the system application discussed in this thesis consisting of many hand held devices and a single terminal device

1.9.1 Banking Authentication

Perhaps the most compelling use scenario for the requirement of a consumer oriented QKD system is to replace the current "chip and pin" infrastructure which is based on the Europay, MasterCard and VISA (EMV) Standard [77]. This system can secure point of sale (POS) and automated teller machine (ATM) transactions and also, alongside another device (such as Barclays' chip authentication program (CAP), known as PINSentry which can be seen in figure 1.26 [78]) can secure remote transactions such as those on the internet.

Figure 1.26: Picture of the Barclays CAP Device, PINSentry

The EMV standards rely on "classical" cryptographic principles [77], which section 1.3 showed do not *guarantee* security. The work described in this thesis proposes a replacement for this specification based on the unconditionally secure technology of quantum cryptography. The technical challenge is to miniaturize the hardware such that the consumer device is of no greater size than the existing CAP readers.

1.9.2 Access Control

Since the user device suggested in section 1.9.1 is necessarily small and low cost, a similar infrastructure could be implemented in any situation where a 100% robust access control system is required. Suggestions for such situations are:

- Physical access control to sensitive areas such as in military or government facilities.
- Electronic access control to any sort of important computer system and/or encryption of secret data.
- Personal access control for personal identification.

1.10 Work Undertaken

Continuing from the progress discussed in section 1.7.2, this thesis details a QKD system suitable for widespread deployment in a consumer setting. The principal novelty in the system is the great discrepancy between the size and cost of the "Alice" and "Bob" devices which is the basis for the applicability in widespread applications. Much of the work undertaken in this thesis has concentrated on miniaturizing the optics of the Alice device and refining various aspects of the system to increase the speed and reliability of key generation.

The work comprises:

- Contributing to the building and testing of a prototype integrated system incorporating repeatable alignment system, on board processing and wireless public channel (Chapter 2 and chapter 5).
- Work towards miniaturizing the optical components required for quantum state generation in the Alice device (Chapter 3).
- Increases to the security and maximum speed of photon detectors in the Bob device (Chapter 4).

1.11 System

The system consists of two major components: A small and portable "Alice" module and a larger, fixed "Bob" module, the internals of which will be discussed in section 2.2. A principal advance from previous work on the system overall is these components' integration into a prototype system demonstrating the suitability for the scenarios depicted in section 1.9.

The "Quantum ATM" shown in figure 1.25 consists of: a 4U 19inch rack mountable case and contains the optics required to measure the quantum states sent from

the Alice devices; a TIA developed in-house at Bristol University by a fellow PhD student, Richard Nock [79], an embedded PC system for data processing and a Wifi router to manage the public channel connections. In order to allow for "plug and play" operation, the front of the Quantum ATM also features a magnetic mount allowing for instant repeatable optical alignment between the Alice and Bob devices.

The "user device" from figure 1.25 consists of: a personal digital assistant (PDA) (although theoretically any pocket device is feasible) which communicates with the Bob device and provides a front-end for the user to interact with; electronic unit for random number generation and limited data processing and an optics unit to generate the quantum states necessary to carry out QKD.

1.12 Summary

This chapter has:

- Introduced concepts of classical cryptography and the basis which historical and modern methods use for security (section 1.2).

- Separately introduced two quantum phenomena important to cryptography; section 1.3 discussed the possibility of a quantum computer being able to break classical cryptography more efficiently than a regular computer and section 1.4 introduced the concept that an arbitrary quantum state cannot be copied since a measurement of the state will disturb it.

- Extended the ideas from section 1.4 into section 1.5 into the concept of quantum key distribution (QKD). A method of generating cryptographic keys which is guaranteed secure by the laws of physics.

- Presented various protocols, attacks on these protocols and improvements to

the protocol chosen for this project (BB84) to improve performance and security. (Sections 1.5.2, 1.5.3, 1.5.3.1 respectively)

- Described analyses of QKD with imperfect devices whilst maintaining unconditional security (section 1.6).

- Provided an overview of various important free space QKD experiments which have been useful to the development of this project (section 1.7.1). Section 1.7.2 then discusses the history of the Bristol system before this work commenced.

- Discussed various commercially available QKD systems in section 1.8 and then in section 1.9 how this work would differ from such systems.

- And finally sections 1.10 and 1.11 describe the work undertaken in this thesis and a brief description of the system as a whole which will be discussed in far greater detail in chapter 2.

Chapter 2

The Bristol QKD System

This chapter consists of an introduction to the actual physical components of the system (section 2.1), the theory of their operation and design considerations necessary for combining them all into a quantum key distribution (QKD) system (section 2.2). Design and construction of the handheld Alice device and integrated Bob terminal was performed during the first year of this PhD and this work culminated in a demonstration at the "SECure COmmunication based on Quantum Cryptography (SECOQC)" conference which is covered in section 2.3. The results obtained for the system exhibited at the SECOQC conference are presented at the end of this chapter.

2.1 System Components

Section 2.2 will provide detailed descriptions of the components in the Alice and Bob devices, a brief introduction to these components is given below and their interconnections depicted schematically in figure 2.1.

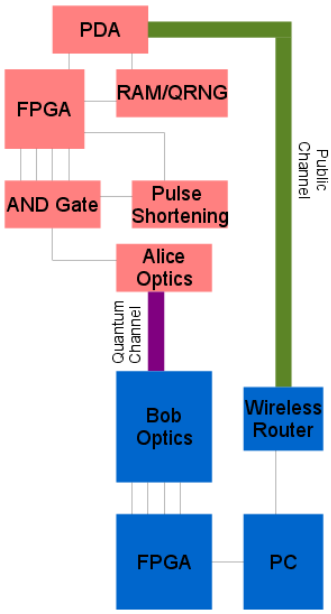


Figure 2.1: Block diagram of whole QKD system showing the Alice components in pink and the Bob components in blue. The lines signify internally connected systems and the Quantum and Public channels (purple and green respectively) signify the system connections outside of the devices. The roles of individual components are mentioned in section 2.1.1 for Alice components and section 2.1.2 for the Bob components.

2.1.1 The Alice Device

PDA This component is the interaction point of the system. A personal digital assistant (PDA) is chosen due to its conformity to the design brief table 1.1 however any device with a USB interface can be utilized here.[1].

RAM/QRNG These components are mentioned together since one of the two will serve as the source of Alice's random bits. For prototyping the same random string can be loaded from random access memory (RAM) repeatedly; this is clearly insecure but isolates the performance of the quantum random number generator (QRNG) from the performance of the system.

FPGA The field programmable gate array (FPGA) is configured to generate some

[1] a laptop is often used for prototyping

clock and select states to transmit based on bits loaded from the RAM or QRNG on every clock cycle. It is the component the PDA instructs to start and stop transmitting.

AND Gates and Pulse Shortening The signals from the FPGA are not sufficiently short to produce photons with accurate enough timing resolution and so a pulse shortening circuit is included. This shortened pulse is then routed to one of the LEDs by way of AND gates. The process of pulse shortening will be discussed in section 2.2.1.1 and is illustrated in figure 2.2.

Alice Optics Briefly, the Alice optics consist of four LEDs, each generating one of the Bennett Brassard 1984 (BB84) polarization states. These are collimated and filtered and sent down the quantum channel. More in depth discussion is in section 2.2.1.

2.1.2 The Bob Device

Bob Optics The optics and various electronics for cooling and quenching the detectors are situated in a light tight milled aluminium box on an xyz kinematic mount to allow for alignment. The optics and such are discussed in much more depth in section 2.2.2.

FPGA The FPGA in Bob is configured to operate as a 4 channel time interval analyser (TIA) which applies a time tag to each detected pulse which is used to reconcile the sparse list of detections Bob generates with the complete list of transmitted states possessed by Alice.

PC A PC motherboard is situated inside the Bob device, the PC interfaces with the FPGA TIA and processes the data along with the subset of Alice's transmitted bits that are sent. Performing all processing on the Bob device allows the Alice

device to be as small as possible, conforming to the design brief (table 1.1)

Wireless Router Wifi is used as a public channel as it does not require line of sight or any further alignment and there is mature and robust hardware and software for interacting over wifi. For the sake of demonstrating the fact that the public channel is untrusted the link is unsecured however this is not necessarily the best option in a real implementation. System administration of the integrated PC is also performed over the network connection.

2.2 Theory of Operation

2.2.1 Alice Optics

2.2.1.1 LED Sources

Single photons are the logical choice for "flying" quantum bits (qubits) in a Quantum Information application; they are easy to produce, have long coherence times and interact weakly with the environment whilst also being easily detectable [80]. Single photons are generated by many physical processes such as Quantum Dots [81], Spontaneous Parametric Down Conversion [82, 83] and Colour Centres in Diamond [84]. It is however, acceptable to heavily attenuate a short pulsed source such as an LED or a Laser providing the statistics are analysed correctly and chosen to minimize signals containing greater than one photon [62]. This is the chosen method in this QKD system since "true" single photon sources are not yet a mature technology and tend to be large, expensive and/or delicate [44] whereas the technology for weakly pulsing an LED or Laser Diode can be done using only basic electronics and neutral density (ND) filters.

There are two options when producing the multiple states required for BB84 QKD, one is to use a single light source and choose the state by employing an

Optical Modulator [85] to select between states, these can, however, be large and expensive. A simpler approach is to use multiple light sources with relevant passive optical elements, this removes the issues of size and cost however adds an extra level of complexity in the requirement that each state must be indistinguishable aside from the property the qubit is encoded in (in this case, polarization).

Another consideration is that in order to produce the short pulses required to ensure a low photon number per pulse, the diodes must be driven with very short electrical pulses, this becomes an issue for Laser Diodes which require a certain threshold current in order to exhibit lasing behaviour, below this the Laser Diode will only undergo spontaneous emission and will behave very similarly to an LED thus invalidating many of its desirable characteristics (yet retaining many of its undesirable ones!). It is possible to construct a current driver to supply sufficient current however this adds further complexity and size to the system whereas an LED can pulse acceptably even when driven with pulses from transistor-transistor logic (TTL) or similar.

Whilst Laser Diodes have many superior qualities compared to light emitting diodes (LEDs), such as narrower optical spectra and very low divergence angles, LEDs are employed in this QKD system for reasons of size, cost and inherent incoherence. Due to the spectral width of the light emitted by LEDs, it is necessary to employ narrow spectral filtering and a great deal of characterization to ensure that no information is carried in the spectra of emitted light pulses.

There is a lower limit on the maximum repetition rate of an LED than of a Laser Diode due to the the differing processes by which the light is produced. An LED is limited by the spontaneous lifetime of the carriers in the LED ($\sim$ns) whereas a Laser Diode driven correctly is limited by the lifetime of the photons in the cavity($\sim$ps). There does lie a potential to outperform the spontaneous lifetime of the LED by using RCLEDs (Resonant Cavity LEDs) [86].

In light of these considerations, a 2×2 grid of LEDs driven by short TTL signals are utilized as photon sources, the pulse width is determined by changing the delay imposed on the second input of the AND gate (figure 2.2), the pulse width W_p is determined by the on time of the input signal T_{on} and the time delay τ_d by:

$$W_p = T_{on} - \tau_d \tag{2.1}$$

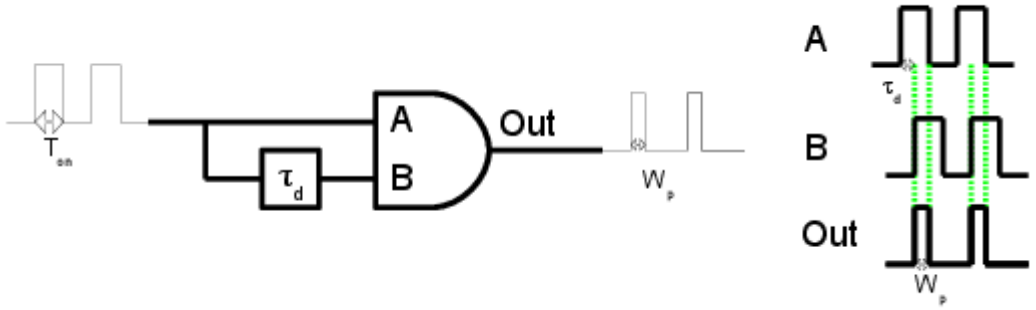


Figure 2.2: Method used to create short pulses to drive the LEDs in the Alice system. A version of the input pulse width T_{on} is delayed by τ_d and recombined with an undelayed version of the input pulse resulting in a shorter pulse of length $T_{on} - \tau_d$ (see equation 2.1

A small piece of polarizing film[2] is placed over each LED aligned in such a way as to polarize the resultant light in each of the four polarizations states depicted in figure 1.5.

2.2.1.2 Collimation Optics

In order to collimate the light from multiple sources, a 2D diffraction grating, hereby referred to as a diffractive optical element (DOE) is employed to combine the light

[2]such as ThorLabs LPVISE2X2

according to the equation [87][3]:

$$n\lambda = dsin\theta_n \tag{2.2}$$

where d is the grating period, n is the diffraction order and θ_n is the angle of diffraction of the nth order.

In the opposite arrangement than if one wished to create a diffraction pattern, the sources are placed at the positions of the first maxima and are therefore combined into a single beam (figure 2.3). It is then necessary to filter out extra diffraction orders (since there are only light sources at the primary maxima) which can be achieved using a pinhole positioned such that it only allows the central peak to pass.

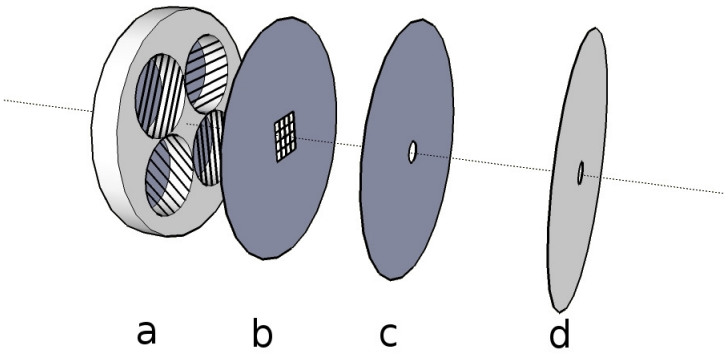


Figure 2.3: The arrangements of the Alice Optics. a) 4 LEDs are secured in a metal housing, each with a polarizer glued across the front corresponding to a polarization required for performing BB84. b) The light from the LEDs is incident onto a diffraction grating which combines the beams into one axial beam (and many higher order terms). c), d) Two pinholes are placed after the diffraction grating to filter out the extra modes from the diffraction grating and then to ensure that the light leaving the final pinhole is transversely coherent. These components are mounted in a 12.5mm optics tube (omitted for clarity).

[3]although this equation is for a 1 dimensional grating, a 2D grating can be considered two 1D gratings placed at right angles

2.2.1.3 Security Considerations

In order to avoid the photon number splitting (PNS) attack(section 1.5.3), the photon number must be chosen such that the count rate is maximised without emitting too many multiphoton pulses. This has been taken into consideration in the security proofs and is a case of calibrating the system to use an optimal μ. In addition to this, aside from polarization, one must ensure that there is no other method for distinguishing photons as this would mean that an Eavesdropper could determine with certainty which polarization was sent.

Photon Number Calibration of the photon number is a surprisingly easy task, one must simply couple a pulsed LED to a single photon avalanche diode (SPAD), correct the count rate for the detector efficiency and then divide the count rate by the frequency of the source.

As previously mentioned, the statistics of the source are assumed to be Poissonian [45] Utilizing the equation: [88]

$$p(N,\mu) = \frac{\mu^N e^{-\mu}}{N!} \tag{2.3}$$

where $p(N,\mu)$ is the probability the source will emit N photons in a pulse of μ mean photon number.

One can say that the probability of a pulse containing a nonzero number of photons is:

$$p\left(N \geq 1\right) = 1 - e^{-\mu} \tag{2.4}$$

where $p(N \geq 1)$ is the probability the source emits more than one photon (ie something that will cause a SPAD to register a count). If the light source is perfectly imaged onto a SPAD, for a sufficiently large data set:

$$p(N \geq 1) \rightarrow \frac{C/\eta}{f} \tag{2.5}$$

where C is the total detected counts, η is the detector efficiency and f is the pulsing frequency. These equations can be combined to obtain an expression for μ for a given count rate.

$$\frac{C}{\eta f} = 1 - e^{-\mu} \tag{2.6}$$

$$\mu = -ln\left(1 - \frac{C}{\eta f}\right) \tag{2.7}$$

In a practical sense, the current can then be varied (using a series resistor) to the LED until the required counts are present on the detector. The general idea is to choose μ such that:

$$p(N \geq 2) = 1 - (1 + \mu)\, e^{-\mu} \tag{2.8}$$

is small enough that Eve's information (she must be assumed to have full information on these bits) can be totally removed in the process of reconciliation.

More specifically, the optimum value for μ can be found by expressing the expected secret rate, R (from equation 1.54) in terms of μ, this equation

depends on several other parameters of the apparatus and so the optimum μ is different for each system[4].

Spatial Information If Eve could simply look into the Alice device and see which LED illuminated for each qubit, she could know with certainty Alice's bit string and therefore send on states to Bob which reflected this, thus gleaming the entire key without alerting either party of her presence. This issue was discussed in section 2.2.1.2 and if the grating and pinhole are selected and positioned correctly, the issue is nullified.

Phase Information If each source emitted light with a characteristic phase, then Eve could utilize some kind of eavesdropping device to measure the phase difference between qubits from Alice and extrapolate this to which light source fired. As mentioned above, this is not an issue with LEDs as they emit incoherent light. It is discussed in [89] that while security can be maintained for a phase coherent source, it is less secure than using a phase randomised source.

Spectral Information Due to the nature of LEDs, they have much wider spectra than Laser Diodes, this presents the possible problem that Eve could measure the wavelength of photons from Alice and discern which source the photon came from therefore being able to forward this state on to Bob in the knowledge that it was faithfully reproduced. A 3nm wide 632.8nm bandpass filter is placed at the exit of Alice's optical path and the power is varied such that the mean photon number across the range of the filter is equal for each LED. Another identical filter at the entrance to Bob's optical path is also implemented although this is useful for reducing the detected background light level and may have benefits in the susceptibility of the system to the back-flash attack mentioned in section 1.5.3.

[4]as discussed in section 1.6.1, this is often close to the channel efficiency

Coherence A critical assumption in the security proofs discussed in section 1.6 is that each pulse from Alice is coherent (i.e. in a single mode; the statement does not imply phase correlations between pulses, which are required to be random). Light emitted from a laser is inherently coherent however the situation is more complicated for a thermal source such as the LEDs used in this apparatus.

As discussed in [73] the longitudinal coherence (coherence time/length) of LEDs is suitable however it is also necessary to consider the transverse coherence properties, the area over which two parts of a beam are coherent with each other. This situation can be visualised by placing a source in front of a double slit and varying the slit separation. For a non point/plane source there is a path difference at each slit introduced due to the non-negligible source size which becomes more important with increasing slit separation until eventually the interference effects stop entirely[5]. The expression for the coherence area diameter, d_c, is [90]:

$$d_c = \frac{0.16\lambda}{sin\frac{1}{2}\theta_s} \tag{2.9}$$

where θ_s is the angular size of the source from the point of view of the observer/interference point and λ is the wavelength of the light.

Therefore a second pinhole must be employed in the collimating optics discussed in section 2.2.1.2. Treating the first (filter) pinhole as a source, and using equation 2.9 one can find the diameter of the coherence area at a certain point away from the first pinhole and employ a pinhole of this diameter to filter out extra modes.

[5]this property is utilized in Michelson's Stellar Interferometer where the light from a star is diffracted by a double slit and the slit separation is varied until interference visibility disappears from which one can infer the angular size of the star far beyond the Rayleigh limit

2.2.2 Bob Optics

2.2.2.1 Beam Splitter Optics

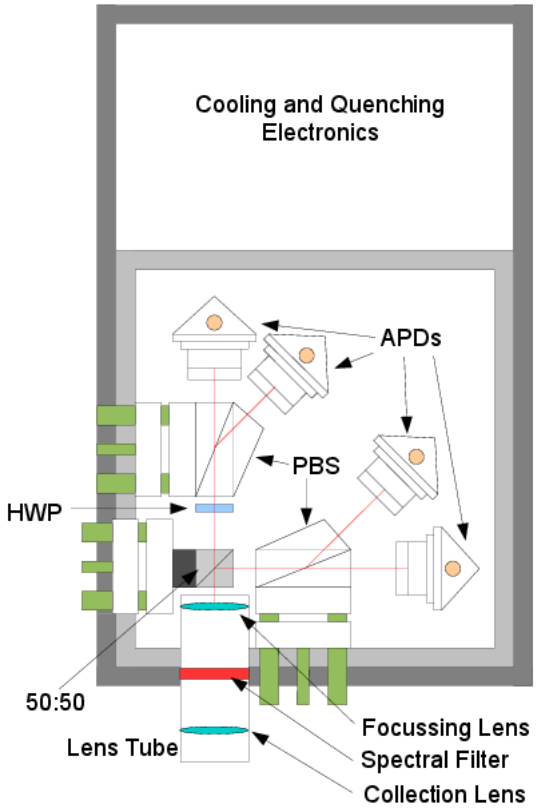


Figure 2.4: Design of Bob Optics box, showing the optical components and the optical path through the device. The light enters the collection lens and is split by the non polarizing beamsplitter (50:50). One path is rotated by 45° by the half wave plate (HWP) and each path is split according to its polarization onto detectors. The beam splitters are placed on tip/tilt stages (green) to allow for adjustments to the alignment.

The optics displayed in figure 2.4 are the standard BB84 polarization discrimination system, utilizing a 50:50 beam splitter which, when single photons are incident upon it, randomly directs the photons along the transmission or reflection arms of the experiment. To allow the optics to be coplanar, a $\frac{\lambda}{2}$ wave plate (HWP) is employed to rotate the polarization of the light in one arm, this is a far simpler prospect

than mounting a polarizing beam splitter (PBS) at 45° to the horizontal. Each PBS then performs the polarization measurement step, transmitting horizontally polarized light and displacing vertically polarized light onto respective detector. The input optical tube on the front of the box contains a two 50mm lenses to image the pinhole on the front of the Alice device onto the detectors and a 3nm wide 632.8nm bandpass filter.

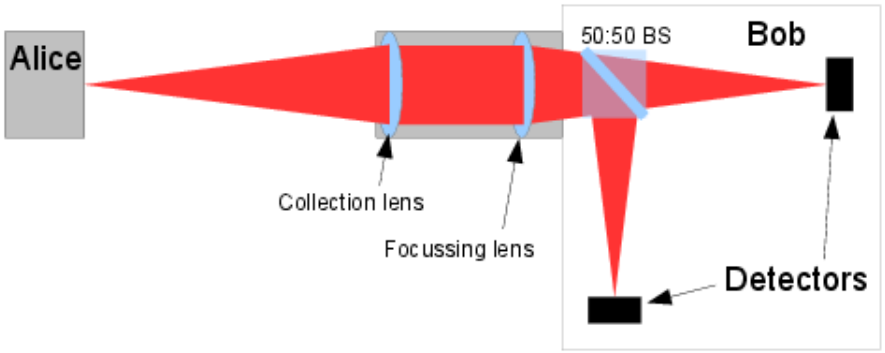


Figure 2.5: Diagram of the imaging system used in Bob (certain elements omitted for clarity). A 10mm diameter 50mm focal length lens focuses on the output pinhole of the Alice device. This beam is then imaged onto the detectors by a second 10mm diameter 50mm lens. The $300\mu m$ pinhole is imaged 1:1 onto the $500\mu m$ detector area allowing for a certain amount of misalignment in the optical path.

Figure 2.5 shows how the pinhole at the front of Alice is imaged onto the detectors in Bob, it is simplified (PBSs removed) for clarity. The optical path length between the imaging lens and the detector is actually longer than 50mm since 35mm of that path occurs in the beam splitters (The 50:50 beam splitters are made from BK7 and the PBS Calcite) resulting in an elongation of the focal length. Utilizing the Optical Path Length equation $OPL = Ln$ [91] where L is the length and n the refractive index. The optical path length of the whole system can be calculated and designed appropriately.

$$OPL_{system} = OPL_{AIR} + OPL_{50:50} + OPL_{PBS} \tag{2.10}$$

$$50 = \frac{L}{n_{air}} + \frac{10}{n_{bk7}} + \frac{25}{n_{calcite}} \tag{2.11}$$

Substituting the dimensions of the components ($L_{50:50} = 10mm$, $L_{PBS} = 25mm$, $n_{BK7} = 1.517$, $n_{calcite,e} = 1.486$, $n_{calcite,o} = 1.658$), the total distance between lens and detector surface of 63.33mm for the o-rays (transmitted) and 61.58mm for the e-rays (reflected).

2.2.2.2 Detectors

The image in figure 2.6 shows the "reach through" design of Avalanche Photodiode utilised in the Perkin Elmer C30902S APDs [92] which are used as photon detectors in the Bob device. The area marked π is the "absorption region"where photons generate electron-hole pairs, the electric field across this area separates the electrons and the holes and moves the electron towards the "multiplication region". In the multiplication region, the electric field is much higher, this is to accelerate the electron to cause it to create more electron-hole pairs via impact ionisation which are accelerated in opposite directions, the holes can then cause further electron-hole pairs further up the potential which causes a self sustaining avalanche breakdown resulting in a high gain [93].

The diodes are reverse biased at 25V above the reverse breakdown voltage, V_B. The breakdown voltage is the point at which the avalanche breakdown process becomes self sustaining and the device is sensitive to single photons (a SPAD). The value of 25V was chosen as the optimum for the device; the efficiency and timing resolution increase with voltage however so do the dark counts, dead time and afterpulsing [95, 96] which are undesirable qualities.

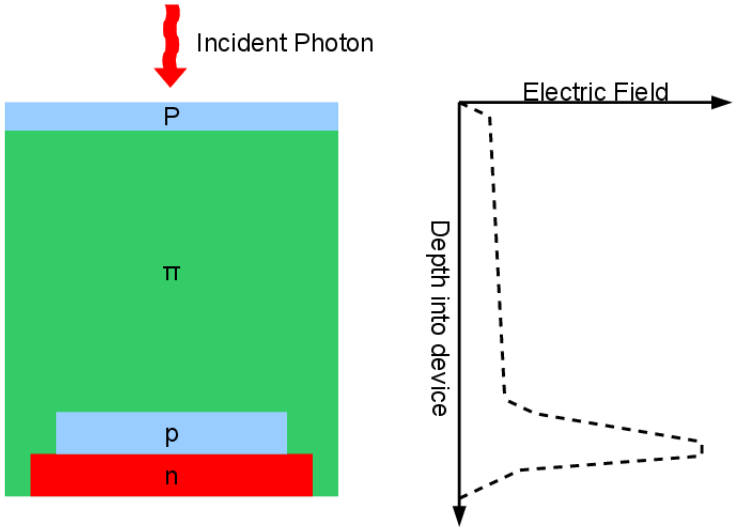


Figure 2.6: Schematic of an APD reproduced from [94]. For SPAD operation, photons enter from above, are absorbed in the p-doped π region and then highly accelerated to create an avalanche of new electron-hole pairs in the high field region on the p-n junction. The graph to the side of the schematic shows the electric field through the depth of the device showing the field across the absorption region to separate the photogenerated carriers and the relatively higher field across the pn junction where the carriers are accelerated such to cause impact ionisation.

If the APD is reverse biased above breakdown, once it begins conducting (due to a photon absorption or a thermal generation of an electron-hole pair) it would continue to conduct and therefore be useless as a time correlated photodetector. If a resistor is placed in series with the diode, as in figure 2.7, before the diode conducts no current is flowing in the circuit so the voltage at A is the same as V_R. When the diode begins to conduct the current flowing through R_L causes the voltage at A to drop and the current to fall below I_{Latch}, the current for which the avalanche is self sustaining. A second resistor, R_S is placed on the other side of the diode and causes the voltage at B to be of the form of the diode current, a spike with an exponential decay [97]. A comparator with an appropriately set threshold can be connected at B (figure 2.8) to provide interface between this circuit and TTL electronics.

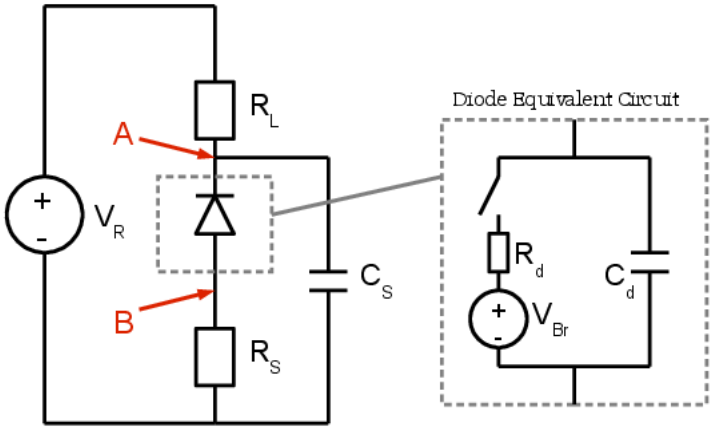


Figure 2.7: The APD in a reverse biased circuit. The diode is reverse biased by V_R through a large load resistor R_L and a small sensing resistor R_S. The stray capacitance R_S is shown and the grey box corresponds to the diode equivalent circuit described in [97] where R_d and C_d are the device resistance and capacitance respectively and the voltage source V_{Br} signifies the breakdown voltage. Points A and B are referred to in the text in the discussion of the passive quenching behaviour.

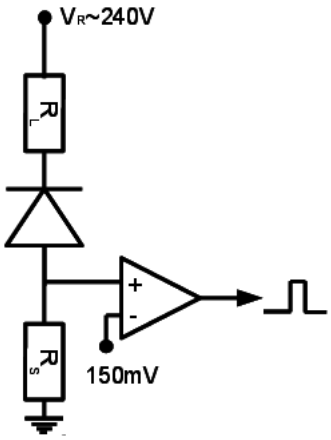


Figure 2.8: A comparator is added at point "B" from figure 2.7, with an adequately set comparison voltage (150mV here), this produces a TTL signal when an avalanche occurs. For adequate quenching $R_L = 470\text{k}\Omega$ and $R_S = 50\Omega$.

2.2.3 Processing

2.2.3.1 TIA

Timing measurement of the photon arrivals is of paramount importance when attempting to establish a shared secret between two parties, in this system it is achieved by an FPGA based Time Interval Analyzer developed at Bristol University [79, 98] which operates based on propagating an incoming pulse down a tapped delay line and comparing each delay tap to the edge of a fast clock (figure 2.9). In this way it is possible to determine the arrival time of a signal to sub-clock accuracy. The FPGA then sends coarse time (clock number), fine time (tap number) and input channel (0-3) via gigabit ethernet to a PC for post-processing.

The diagram in figure 2.9 shows a schematic of the TIA operation. Time tags are generated by comparing variously delayed versions of an input pulse against a clock signal. In this way the pulse can be identified by its coarse arrival time and also the number of delays it had passed through giving a timing resolution of the clock period divided by the number of delay taps. In reality the gates are more complicated than this however the detailed description of this device is not the purpose of this document and further information can be found in [79]. The actual device has a clock speed of 225MHz and 146 delay taps allowing a coarse time resolution of 4.4ns and a fine resolution of 30.5ps. Tests performed in [79] show an RMS timing measurement error of $\sim$ 35.2ps.

2.2.3.2 Algorithms and Reconciliation

Once the data is accrued on the PC, since no clock is exchanged between the Alice and Bob devices, the data must be analysed to discriminate between signal pulses and noise pulses, this is done by "gating" the data. Gating is performed by Alice sharing the clock *rate* but no other clock information, with this, an arbitrary clock

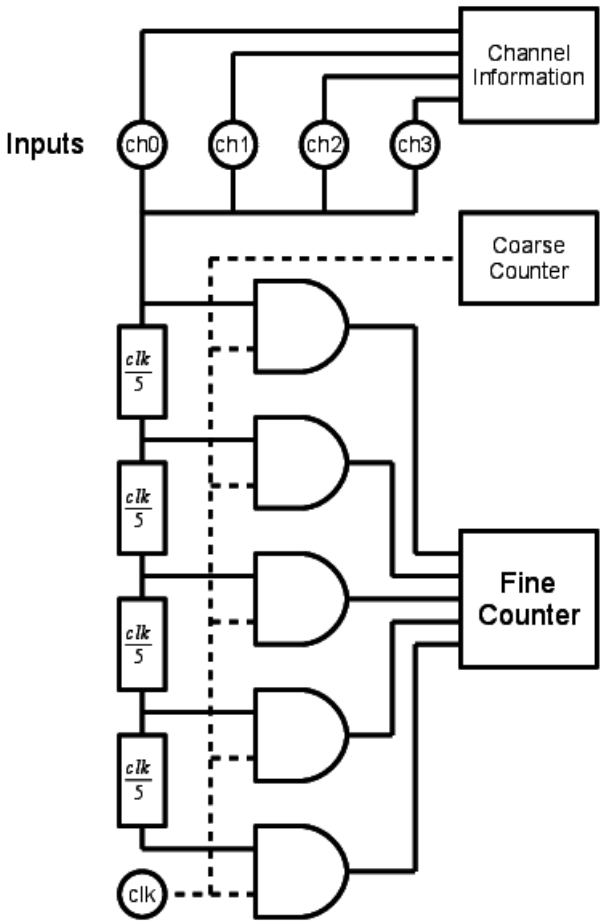
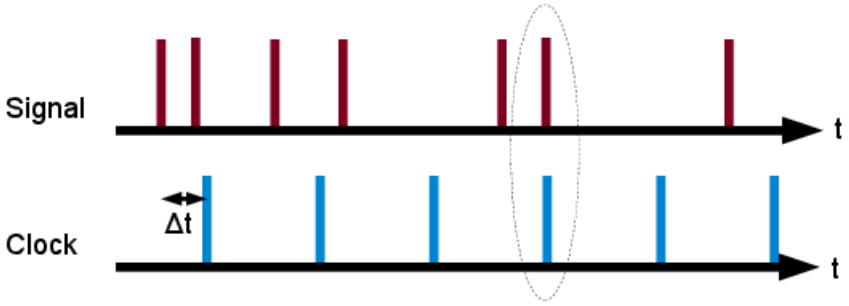


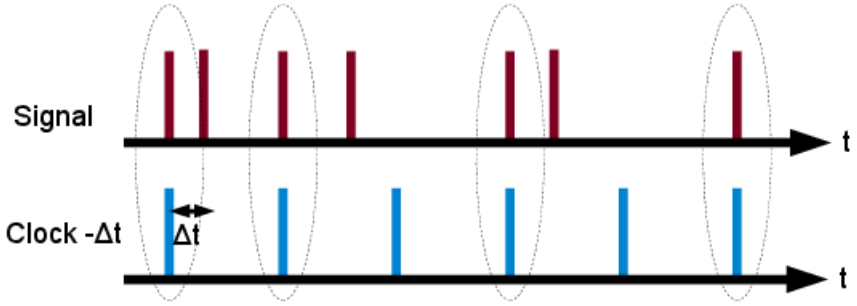
Figure 2.9: A schematic to demonstrate the operation of the TIA utilized in the system. The circles "chX" are the inputs; the blue boxes correspond to electronic delays; the orange, AND gates and the green to information outputs. Refer to section 2.2.3.1 for details of the operation.

can be simulated against which the time of arrival of all signals can be fitted (figure 2.10); anything not congruent ($\pm$ some width $\approx$ the timing uncertainty) with the gated data can be considered background noise.

Data is then shared with Alice as per 1.5.2.1 via a classical channel, in this case, a wifi network, the errors are removed via the CASCADE error correcting protocol and then a privacy amplification algorithm is run in order to minimize the knowledge gained by Eve from errors in the system. More on these processes is discussed in chapter 5.



(a) Comparing the detection times with the clock. One detection is found congruent with the clock.



(b) The clock times are shifted by Δt. There are now four detections congruent with the clock

Figure 2.10: The process of comparing the detection times with a variety of clock delays to isolate the signal from the noise.

2.3 SECOQC Demonstration System

The system inherited at the beginning of this project functioned correctly however it consisted of two similarly sized devices aligned on an optical breadboard with external connections to desktop PCs for all processing (figure 2.13(a)). This proved principle for the idea of a consumer grade QKD system but did not really provide practicality as discussed in table 1.1.

Aside from packaging the devices into suitable enclosures, the Bob optics were redesigned to include simpler tuning of alignment and the issue of aligning the two devices was solved by the use of a magnetic kinematic mount.

Following the events in this chapter, the system was largely dismantled due to some damage incurred in transport. The opportunity was taken to isolate the subsystems and improve them individually. For this reason the results are presented separately from the system performance characterisation performed in chapter 5.

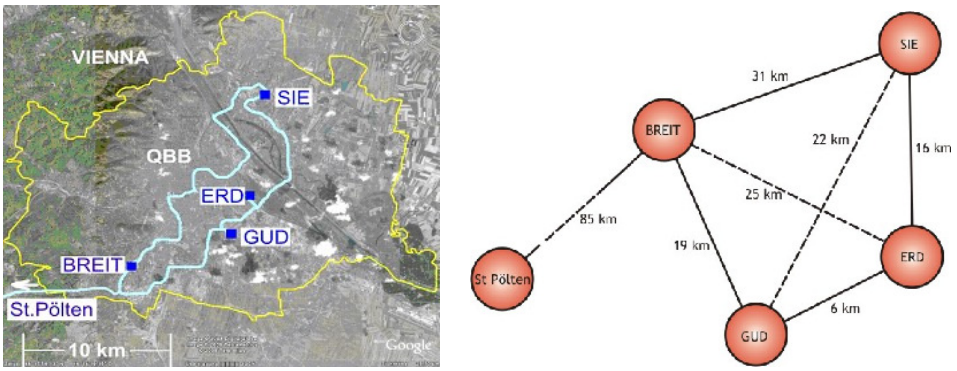
2.3.1 The Event

This project was demonstrated at the SECOQC Conference between 8-10th October 2008. This conference was the first demonstration of a working QKD network [99] and featured a "Quantum-Back-Bone" network comprising of five nodes (figure 2.11), four of which were in Vienna and one 85km distant in the nearby city of St. Pölten. In all 7 different technologies were demonstrated :

- Auto-compensating Plug&Play [74]
- Coherent One Way [100]
- One Way Weak Pulse [101]
- Entangled Photons [102]
- Continuous Variables [103]

- Last Mile Free Space QAN [104]
- Low Cost Secure Key Exchange (the subject of this thesis)

It is important for the widespread adoption of QKD that standards exist to allow for the inter-operation of different devices and for simplicity of maintenance once large networks are deployed. Rigorous standards also allow for stringent bounds on system performance allowing peace of mind that security is guaranteed. In order to realise this the SECure COmmunication based on Quantum Cryptography (SECOQC) project, along with ETSI, the European Telecommunications Standards Institute, formed the Quantum Industry Specification Group, QISG [105].



(a) A Map of the SECOQC Network (b) A Diagram of the SECOQC Network

Figure 2.11: The SECOQC Network in Vienna

2.3.2 System

Previous to the demonstration in the SECOQC project, the physical system was developed on a large optical bench with a separate PC both sides for the classical processing [72] (figure 2.13(a)). This was sufficient for demonstrating the quantum channel link however once this was proved viable it was necessary to develop the points discussed in table 1.1 such that the model of a trivially replaceable Alice module could be demonstrated. The system devised for use is schematically displayed in figure 2.12 and a photograph is shown in figure 2.13(b).

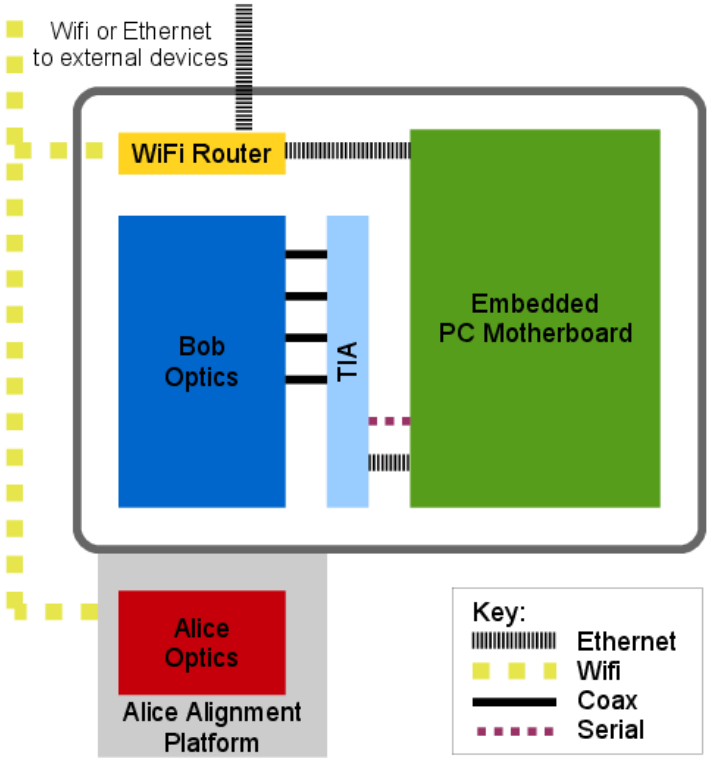
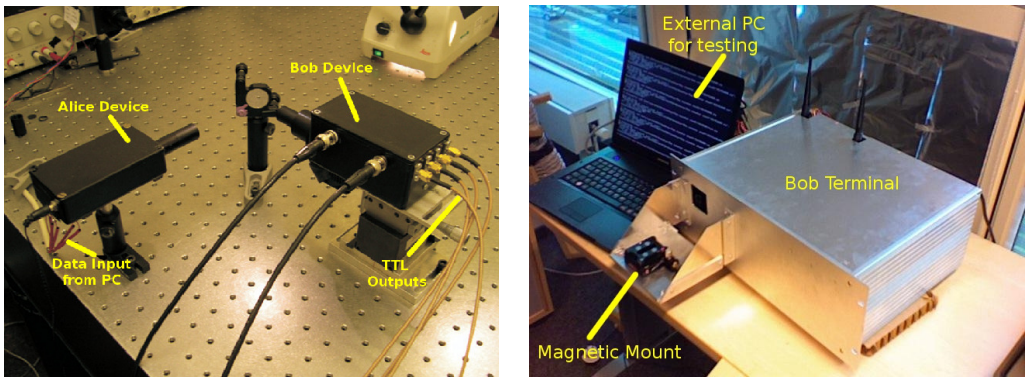


Figure 2.12: The components included in the QKD Terminal device demonstrated at SECOQC Conference

In order to demonstrate the possibility of developing a hand held Alice device, a circuit was developed with drive electronics and was controlled using a PDA (HP iPAQ) via serial expansion card. Clearly this is not a practical system for a production device however many of the components used on the drive electronics are duplicated in modern mobile devices and this application could be incorporated into their design at a far lower level. The electronics were placed into an aluminium box fitted with 12.5mm optics tube containing the LEDs, DOE and pinholes.

The rest of the device is referred to as "The Bob Terminal" and consists of a 4U 19-in rack steel casing (the dark grey curved edged rectangle in figure 2.12) with a platform bolted to the front (light grey rectangle) for the Alice mechanism. A photograph is displayed in figure 2.13(b)



(a) Old System (Optics only) showing Alice, left and Bob, right.

(b) SECOQC Demonstration Apparatus (Bob only, Alice removed so magnetic mount is visible)

Figure 2.13: A Comparison between the system as of the 2006 publication [72] and the SECOQC Demonstration

For Alice alignment, a magnetic kinematic mount base (Newport Optics BK-2A) was bolted to a tip/tilt stage (Thorlabs APY001) for two of the degrees of freedom of alignment required to get best coupling of light from Alice to Bob. The magnetic mount allows better than 100microradians pointing accuracy [106] for repeated replacement onto the mount, corresponding with a $\pm 15\mu m$ error on the position of the spot image on the detector. Since a $300\mu m$ pinhole is being imaged onto a $500\mu m$ detector active area, this ensures adequate alignment and no significant decrease in coupling is expected with repeated placements.

The Bob optics box is mounted on a 3 axis (XYZ) translation stage (Newport Optics ULTRAlign 562) bolted to the bottom of the case providing the remaining degrees of freedom (aside from rotation) required for alignment.

The TTL output from the comparators in the detector circuit in section 2.2.2.2 is fed directly via SMA Coaxial cables into the four inputs on the FPGA TIA board. The TIA is connected via two separate connections to the PC: serial, for control and gigabit ethernet, for data. The PC is a "headless" mini-itx motherboard. Communication is made with this device over secure shell (SSH) or virtual network computing (VNC) via the router, either over the wifi network or through an ethernet

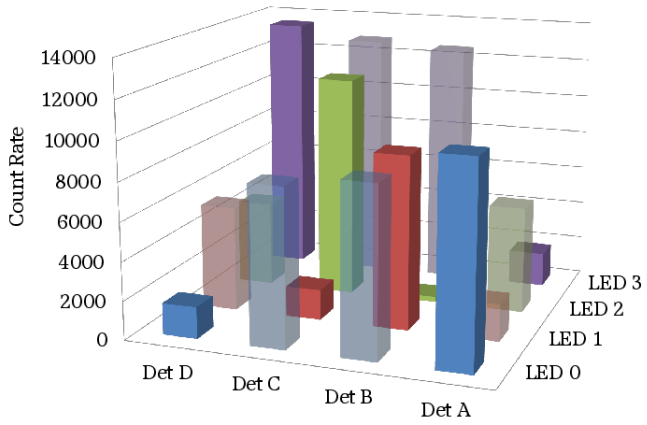
port situated on the rear of the case.

Unsecured wifi was chosen for the public channel as it is a mature and increasingly ubiquitous technology with high bandwidth and is easily "hackable" in that it inherently proves the quantumness of the key distribution as any arbitrary eavesdropper could be invited to utilize one of many available packet sniffing techniques to intercept the classical communication in full [107, 108]. In practice any other classical communication protocol could be chosen (bluetooth, radio, IR) to fit later device specifications.

2.3.3 Results

In order to characterise the system, the extinction ratios of the detectors need to be determined. The extinction ratio is defined by the proportion of light sent in one state (say, 0) and detected in the opposite (in this case, 1). If the count rates on each detector are plotted for each LED illumination, each row and column should have one element containing 50% of the counts, two containing 25% of the counts (these correlate to the opposing basis) and one element containing 0%. This, of course is not the case, and the proportion of counts in the 0% element to the total counts in the basis is called the extinction ratio. The data taken in the Vienna System are shown in figure 2.14.

The repeatability of the magnetic mount operation was then tested by repeatedly replacing the Alice device into the magnetic cradle and measuring the count rate and average extinction ratio; this data is displayed in figure 2.15 and shows that there is no significant down shift to misalignment. More rigorous testing of this method has been performed since the SECOQC conference and is presented in chapter 5.



(a)

LED-Detector Pair	Extinction Ratio
0A	9.50%
1B	4.03%
2C	7.14%
3D	6.80%
Average	6.87%

(b)

Figure 2.14: (a)Matrix of count rates across all 4 detectors for each LED illuminated. The bars corresponding to the non signal basis are dimmed out for clarity. (b) a table of the extinction ratios, the proportion of opposite signals detected for a given transmitted bit

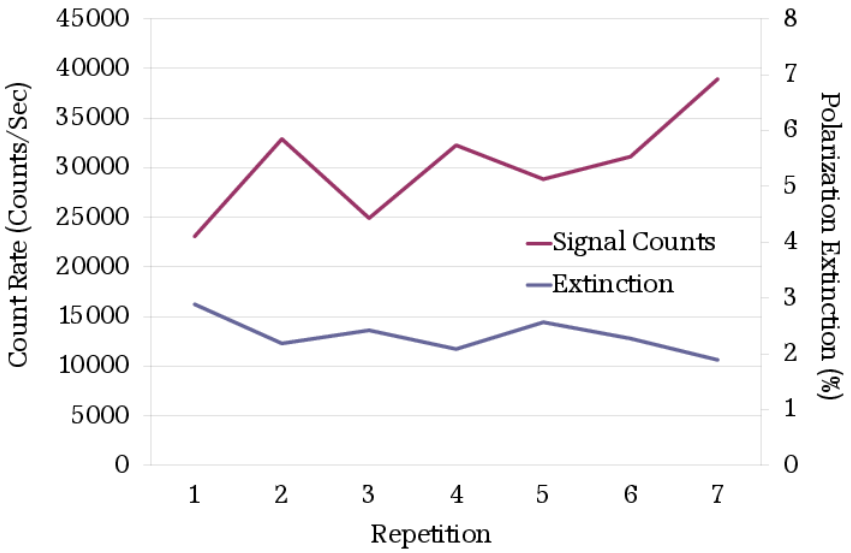
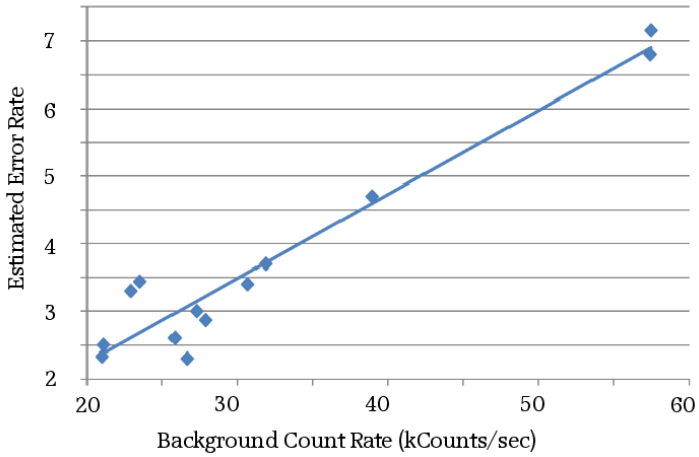


Figure 2.15: The Alice device was repeatedly re-docked on the magnetic mount and the total counts and Polarization Extinction Ratio in the "1B" basis were measured.

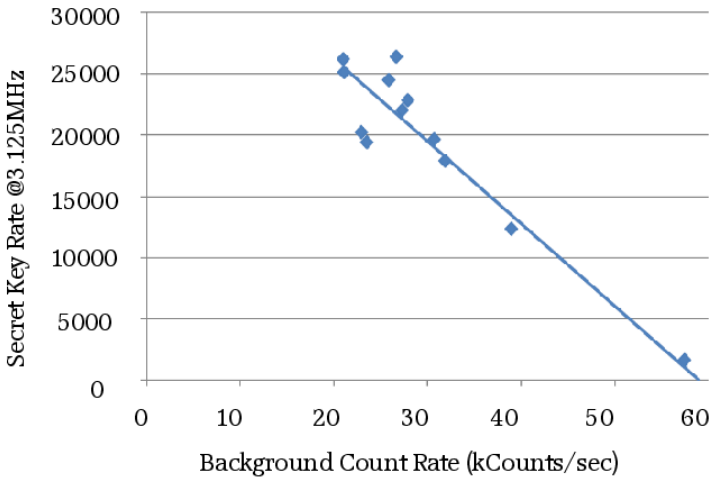
Raw key was generated with a range of different background light levels. A graph of the quantum bit error rate (QBER) of the raw key and corresponding secret rate are shown in figure 2.16. A background rate of 40-45kCounts/Sec was measured in an indoor windowed room in daytime verifying that the system would produce an estimated rate of 10kBits/Sec of secret key in general daylight operation. The parameters of the system are detailed in table 2.1.

Repetition Rate	3.125MHz
Source Wavelength	632.8nm
System Efficiency	10%
Mean Photon Number	0.3
Transmission Time	4 seconds
Gate Width	5ns

Table 2.1: Key parameters of the system used in the SECOQC demonstration.



(a)



(b)

Figure 2.16: QBER (a), and corresponding key rate (b) for varying levels of background for system demonstrated in SECOQC conference. The key rate was calculated using the GLLP [62] proof which was discussed in section 1.6.1

2.4 Summary

This chapter has:

- Provided a description of the Bristol Low Cost Short Range QKD system.
- Discussed theory behind some of the design aspects of the system.
- Addressed security considerations of the components used.
- Given a summary of the SECOQC conference and the performance of the system as of that time. This was the culmination of the first year of this PhD.
- Shown key generation of $\approx$ 10kB/sec in indoor daylight with a magnetic docking method for alignment.

Chapter 3

The Alice Device

The Alice device consists of the optics necessary to create photons in the four polarization states necessary for Bennett Brassard 1984 (BB84) and relevant electronics for converting a random string into driving pulses for the optics. The optics comprise a 2x2 array of red LEDs with a polarizing film placed over each of them. The light from the LEDs is combined using a 2D diffraction grating and then spatially filtered using pinholes. A field programmable gate array (FPGA) is used for control and to interface with some further device (PDA, laptop, mobile phone).

The work undertaken in this chapter focuses on research into methods for decreasing the physical size of the optics module by concentrating on the processes of producing polarized light and of collimating the separate sources into one.

Figure 3.1(a) shows the current state of the Alice device, as it stands it is clearly unsuitable for consumer applications, it is large and somewhat expensive. This is not cause for concern however, figure 3.1(b)[1] shows a picture of a prototype iphone from approximately 2005, it measured 5x7x9 inches, clearly not suitable for usage as a mobile phone. The relevance of this is that it is not unreasonable to expect that once the specifications of the Alice device have been settled upon and enacted,

[1] `http://arstechnica.com/apple/2013/03/exclusive-super-early-iphone-prototype-had-5x7-screen-serial-port/`

(a) The Alice device (b) Early iPhone prototype

Figure 3.1: Comparison of the current Alice device to a recently unearthed picture of an iPhone development prototype.

then the degree of miniaturization that got figure 3.1(b) to a genuinely pocket size device could be achieved with the device in figure 3.1(a)

Figure 3.2 shows an idea of what a low cost device might look like, this is in no way a fixed goal however it is useful to consider in order to give direction to the research into improving the system. Another similar idea for a final product would be to produce an off-the-shelf integrated circuit (IC) which would be able to be put into any generic device for quantum key distribution (QKD) in a similar manner to how near field communication (NFC) chips are becoming commonplace in smart phones.

The goal of a credit card style device shown in figure 3.2, the IC idea mentioned above and the mobile phone/personal digital assistant (PDA) type device shown in figure 1.25 all have sufficient in common at the moment to derive some general directions for research to progress in.

In addition to the device centric requirements, there are general requirements such that the transmission speed should be as fast as possible and not compromise the security of QKD and probably most importantly that it is as convenient to use as the current security methods it seeks to replace (for now we will consider PINSentry, referred to in 1.9.1).

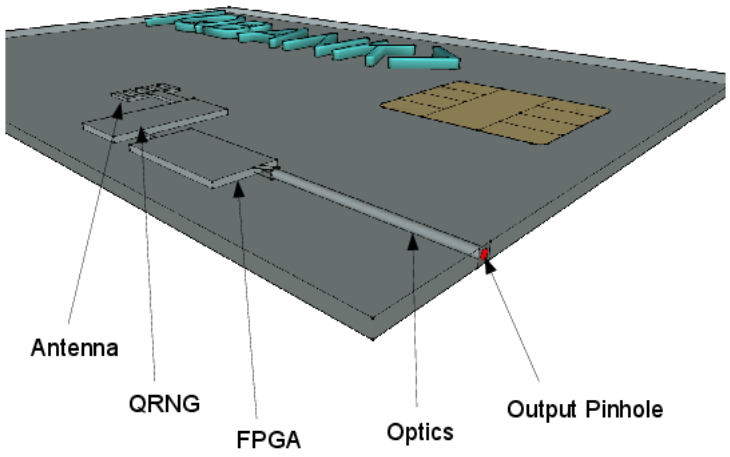


Figure 3.2: Artist impression of a possible final Alice device wherein the components are minimized to the degree that they could fit into a credit card style device.

This chapter will introduce some ideas which will allow for greater miniaturization or faster transmission and will begin to investigate how to enact those ideas.

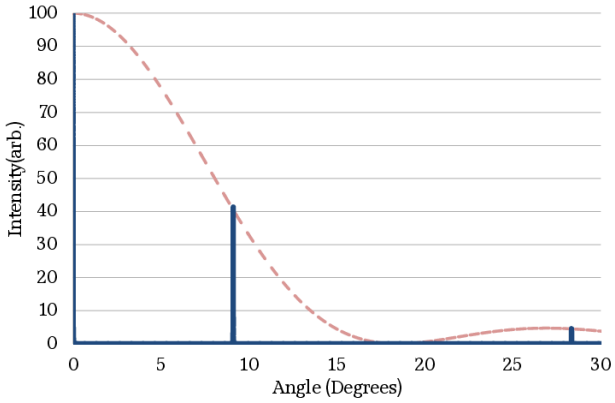
3.1 Ideas

3.1.1 Remove need for diffractive optical element (DOE)

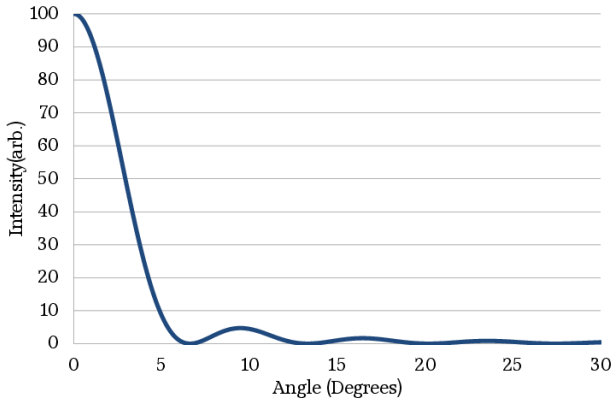
The use of a diffraction grating to combine the light from Alice is fundamentally a good one, due to the high amplitude first order diffraction terms of a grating there can be high throughput and good indistinguishability. One drawback however is that as well as the x,y alignment between the centre of the grating and the LEDs, rotation is of concern as the diffraction peaks are points and a rotation could lead to a greatly decreased transmission as the peak is moved from the LED.

A far simpler idea is to use a pinhole to combine the light [109], a pinhole has infinite rotational symmetry so there is no angular dependence on the diffraction and alignment is greatly simplified. It is necessary to analyse both systems to determine

whether the decreased transmissivity of a pinhole based beam combiner is within acceptable bounds or not.



(a) Diffraction from a grating



(b) Diffraction from a pinhole

Figure 3.3: Comparison between the 1D diffraction from (a): a grating and (b): a pinhole. The dotted red line in (a) is the single slit "envelope" resulting from the number of slits illuminated being finite.

Figure 3.3 compares the diffraction patterns for a grating and a slit, the dimensions of the grating are those of the grating currently used in the Alice collimator optics (750 2μm slits, 4μm separation) and a pinhole that produces a first maximum at a similar angle (5.5μm). It is clear that the grating produces a much higher and narrower first maximum than the slit, the area under the first maxima for each pattern was calculated and is displayed in table 3.1.

Proportion of light in the first maxima	Percentage
Diffraction Grating	39.14%
Single Slit	4.81%

Table 3.1: Percentage of light contained within first diffraction peak of Single Slit and Grating

To rigorously analyse the optical arrangements, it will be first necessary to define some measurements in the system and then some measures of quality. The measurements are shown in figure 3.4, as well as this, a quantity henceforth known as the "lumped efficiency" or ϵ will be defined as the proportion of light emitted by an LED which will exit the second pinhole.

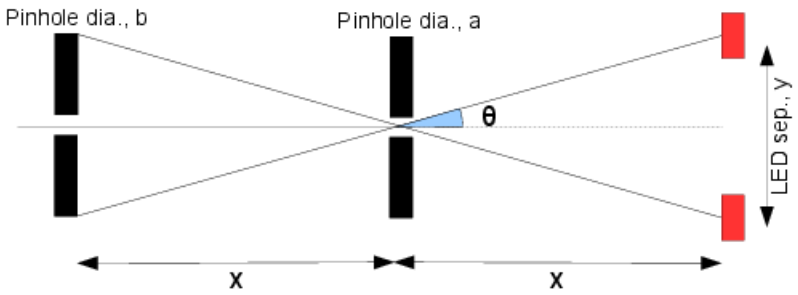


Figure 3.4: Schematic of the simplified system wherein a pinhole is used to collimate several beams of light.

For a given first pinhole size, a, it is first necessary to calculate where the diffraction maxima are situated, this is achieved by use of the Fraunhofer Single Slit equation [87]:

$$I(\theta) = I_0 \frac{\sin^2\left(\frac{\delta}{2}\right)}{\left(\frac{\delta}{2}\right)^2} \tag{3.1}$$

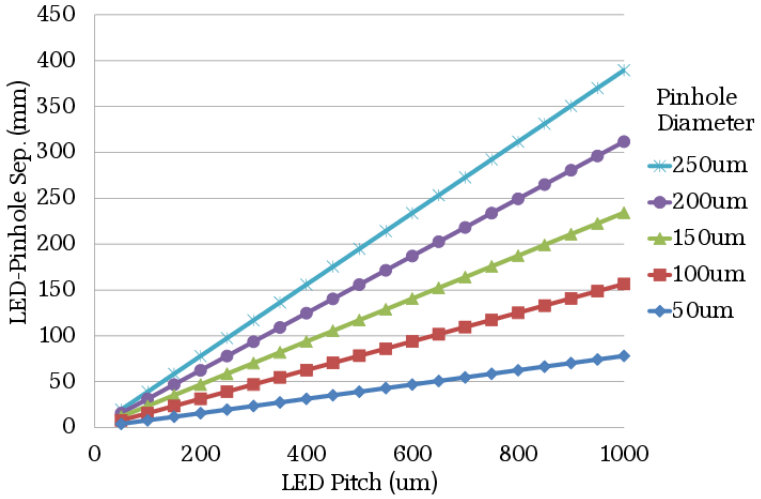
where

$$\delta = \frac{2\pi a \sin\theta}{\lambda} \tag{3.2}$$

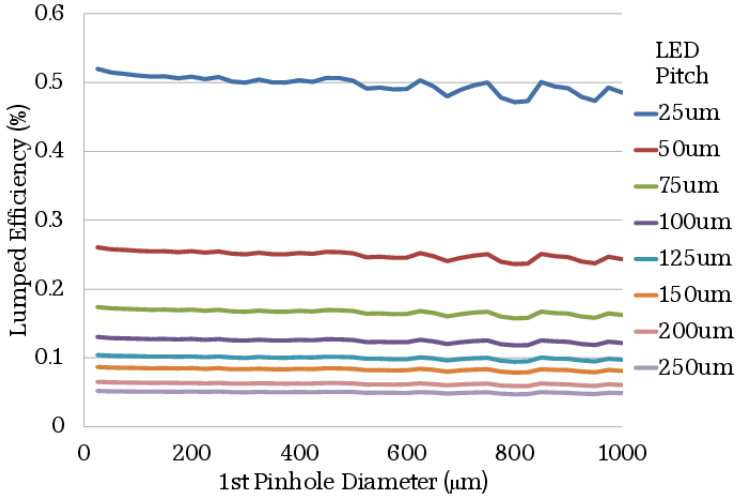
Once the angle of the first maxima is determined, application of trigonometry can be used to determine the LED-Pinhole separation, x for a given LED pitch and pinhole diameter. The transmitted light from pinhole A is then required to be spatially filtered to block out all non axial modes. This can be done by determining the angular width of the first maxima of the diffracted light from Pinhole A and placing a further pinhole, B in a position such that any light transmitted outside of the cone defined by the first minima is blocked.

From figure 3.5(a) it can be seen that even for very small (sub millimetre) scale pinholes and LED pitch values, the LED-Pinhole separation can still be very large. Since in table 1.1 we established that size is of great importance to the specifications of the system, it is therefore necessary to choose pinholes and LED pitches as small as is practical to reduce the length as much as possible.

It is necessary, however, to consider the impact that the choices made in light of figure 3.5(a) with regards to the lumped efficiency of the optics. As can be seen in figure 3.5(b), the pinhole diameter has negligible effect compared to the LED Pitch which should be selected as small as possible. The effects of significantly lowering the lumped efficiency are discussed in section 3.1.3.



(a) LED Pinhole Separation for Various LED array pitch and pinhole values.

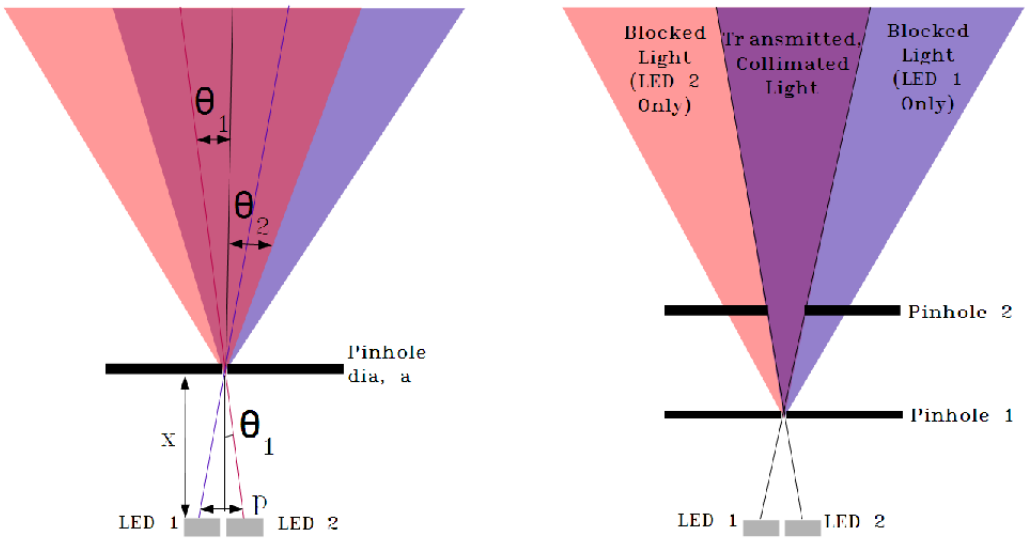


(b) Throughput of the Pinhole system for various pinhole diameters and LED array pitches.

Figure 3.5: Characterisation of the Pinhole collimating scheme simulated with a FWHM LED emission angle of 25°. Note how the lumped efficiency (throughput) depends only very weakly on the pinhole diameter. A far more important quantity in the design is the LED array pitch which should be as small as possible to minimize the physical size of the whole system and maximize the throughput.

3.1.2 Diffractive Overlap

A further possibility, which would be possible only with very small LED pitches ($\sim 100\mu m$), is to utilize an arrangement wherein the zeroth maxima from each LED being diffracted through the pinhole have a certain degree of overlap (figure 3.6(b)), in a similar manner to the previous system a secondary pinhole could be employed to spatially filter light in non-overlapping regions. Felicitous choice of pinhole diameters can lead to higher lumped efficiencies than in the previously described method which can be seen from comparing figure 3.5(b) with figure 3.7.



(a) (b)

Figure 3.6: Diagram of the principle behind collimating the zeroth diffracted peaks from a pinhole.(a) demonstrating the diffraction angles defined for the simulation parameters and (b) shows how a second pinhole filters only for the overlapping region. (Colour is used to clarify the mixing of light from LEDs rather than any kind of spectral information)

To investigate the effect of pinhole diameter and LED pitch on the lump efficiencies, a simulation was performed, again utilizing the Fraunhofer diffraction equation (equation 3.1) and analysing how much of the light incident at an angle θ_1 crosses the pinhole axis (figure 3.6(a)), represented by the angle θ_2. Due to the symmetry

of the system the overlap segment can then be defined as the angles between θ_2 and $-\theta_2$ from the pinhole axis or $\theta_2 + \theta_1$ and $\theta_2 - \theta_1$ from the LED axis.

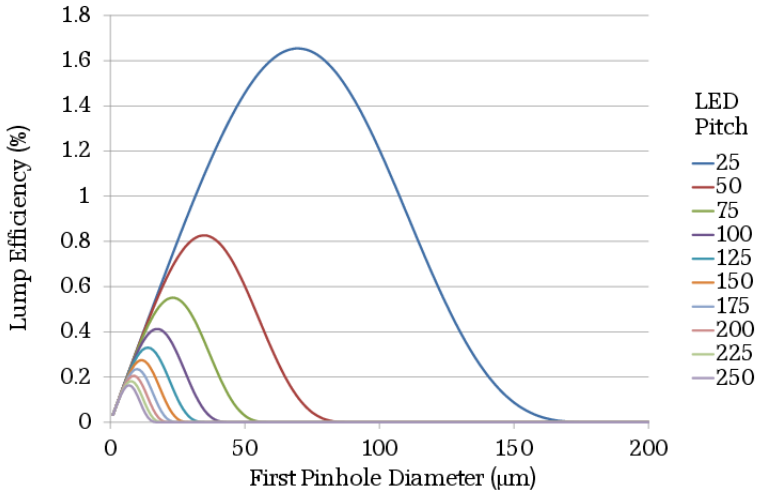


Figure 3.7: Efficiency of the diffractive overlap system for various parameters of the arrangement.

The proportion of the light in this segment of the diffraction pattern compared to the total under the diffraction pattern from $\frac{-\pi}{2}$ to $\frac{\pi}{2}$ constitutes the *diffractive throughput* of the system. Combining this with the amount of the LED emission (estimated as Gaussian with a 30° FWHM) collected by the diffraction pinhole in the first place gives the lump efficiency which was calculated for a range of pinhole diameters over a range of LED pitches (figure 3.7)

The curves in figure 3.7 show an optimum value for pinhole diameter for a given LED pitch, this is due to larger pinholes diffracting less (smaller overlap areas) but allowing more light to pass through them. The trend for higher lump efficiencies with smaller LED pitch can be clearly understood since this is a reduction in θ_1. Comparing this data to the data in figure 3.5(b), the same trend for increasing efficiency at smaller pitch can be seen although the absolute values for the overlap method are higher if the dimensions are chosen correctly. In light of this it is this system which will be further analysed in the following section.

3.1.2.1 Coherence

As per section 2.2.1.3, a further pinhole is now necessary to restrict the light to a single mode. Unfortunately, this will reduce the lump efficiency even further by introducing a further filtering step. All is not lost, however since there is a second pinhole present anyway its size can simply be reduced if it is bigger than the coherence area of the diffracted light.

Recall, equation 2.9:

$$d_c = \frac{0.16\lambda}{sin\frac{1}{2}\theta_s} \tag{3.3}$$

d_c can be converted to a "coherence half angle" θ_C by simple trigonometry once an appropriate pinhole-pinhole distance is chosen. Assuming θ_C is smaller than θ_2, a pinhole of diameter d_c can then be placed to ensure single mode output of the arrangement. Since θ_C is independent of LED pitch, a graph of first pinhole diameter vs d_c is shown in figure 3.8 and the consequent reduction in lump efficiency is graphed in figure 3.9.

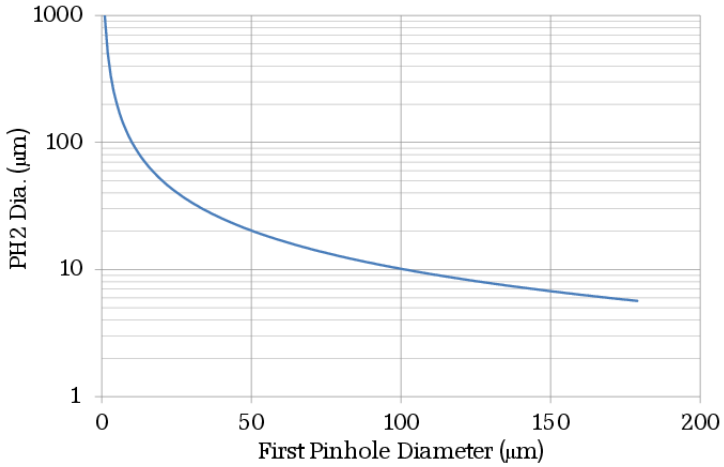


Figure 3.8: Second pinhole diameter (equal to the coherence area) for varying values of First Pinhole diameter at an example fixed distance of 10mm.

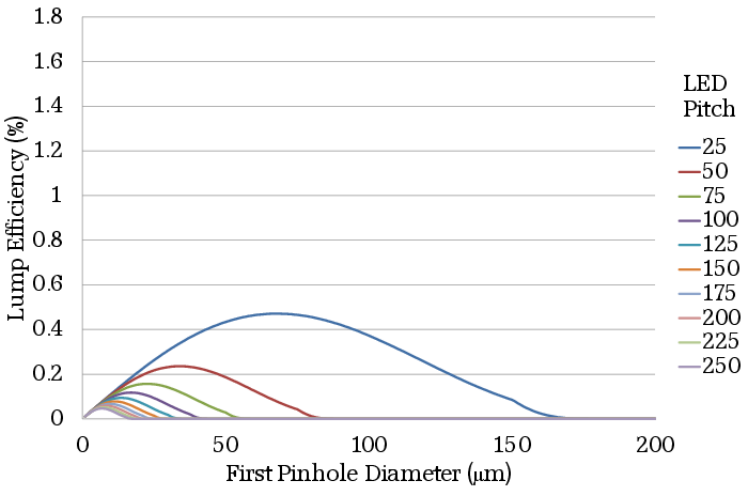


Figure 3.9: Same simulation as figure 3.7 but for a choice of Second Pinhole which is smaller than the coherence area, ensuring single mode operation. Note the scale is the same as figure 3.7 which highlights the drop in throughput.

3.1.3 Collimators Summary

It should be noted here that although the lumped efficiencies discussed in this section are very low ($< 1\%$) this should not matter. A sample LED (HLMA-QH00 [110]) was measured to have a maximum CW output power of $1.32 \pm 0.01 mW$. Pessimistically bounding the pulse width at 1ns, at 10MHz gives 99% attenuation in pulsed mode and a further 1% attenuation from the lump efficiency yields $1.32 \times 10^{-7} W$. Comparing this to the required power which is calculated by multiplying the single photon energy $= \frac{hc}{\lambda}$ with the average photon number per pulse (0.1) and the repetition rate (10MHz) which is $3.14 \times 10^{-13} W$ means that the 1% throughput is clearly acceptable as long as sufficient current is supplied to the LED.

The optical arrangement presented in section 3.1.2 is recommended to be utilized in the next generation Alice device if sufficiently small LEDs on a sufficiently close pitch can be obtained and an effective way of mounting polarizers to these is found.

3.2 Micro Polarizers

Currently, the polarization states necessary for BB84 are generated by separate LEDs each with polarizing film placed over them in the required alignments. This is practical for LEDs of large size ($\sim 1mm$) however the practicality of aligning polarizing film to sub millimetre scale devices is challenging and a more advanced technique is necessary.

Visible optical polarizers, such as the film mentioned above, are generally made from micro or nano scale particles upon a film or glass substrate. Typical materials are crystals of iodoquinine sulfate [111] or metal nanoparticles [112]. A practice that has been used extensively for longer wavelengths [113, 114] is that of the *wire grid polarizer (WGP)*. The WGP works by presenting an array of parallel wires perpendicular to an electromagnetic wave (figure 3.10). The components of the electric field of this wave that are parallel to these wires couple strongly to the electrons in the wire resulting in absorption or reflection. The electrons excited by the components of the E field perpendicular to the wires are not so free to move and the wave passes through the grid.

The criteria for a WGP to be effective in polarizing the light is that the grid pitch and wire width (b and d in figure 3.11 respectively) must both be significantly smaller than the wavelength of the light in question [115]. For the case of this QKD experiment, the light is of wavelength 633nm. Recent advances in nanofabrication have made it a viable possibility to experiment with wire grid polarizers for visible wavelengths. As can be seen in [116], the polarization extinction ratio (PER)(see equation 3.4) decays rapidly with increasing grating pitch.

$$PER = \frac{P_{TE}}{P_T M} \tag{3.4}$$

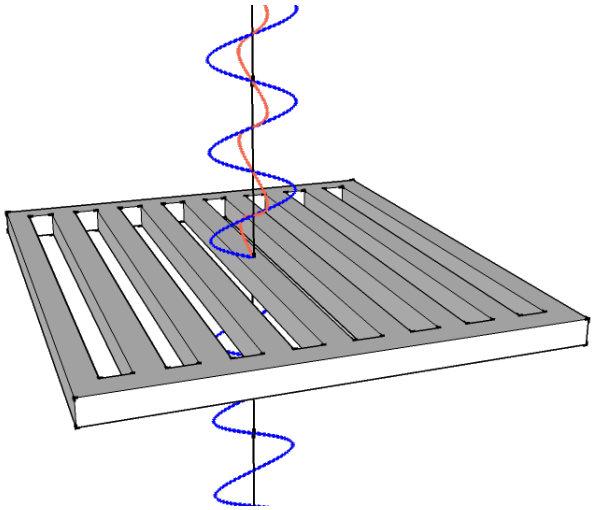


Figure 3.10: A Wire Grid Polarizer and how it interacts with the E field of an EM wave. E waves parallel to the wires induce currents along the length the wires which dissipates the energy of the waves; E waves perpendicular to the wires the induced currents are across the width of the wires which provides less freedom for electrons to move thus dissipating less energy.

where P_TE and P_TM are the power transmitted in TE and TM respectively.

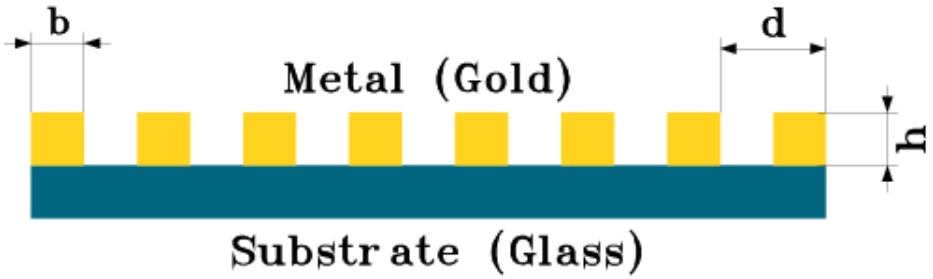


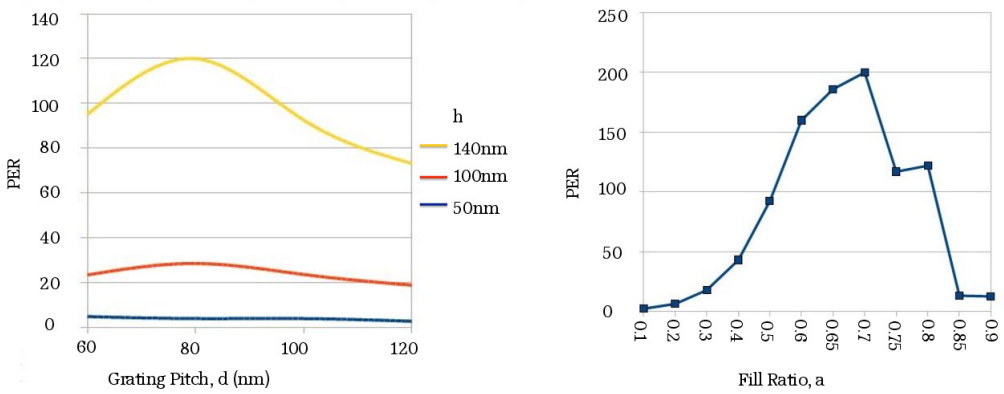
Figure 3.11: The quantities varied during modelling. Not shown is a, the "fill ratio" defined as $\frac{b}{d}$

As a proof of concept, FDTD simulation was performed on a range of subwavelength gratings the results being shown in figure 3.12. Figure 3.12(a) shows how the PER, varies with the grating pitch for a variety of metal thicknesses. This shows that a thicker wire will produce better PER, which is expected due to the the higher chance of absorption of light polarized parallel to the wires. This simulation also indicates that the optimal pitch for 633nm light is 80nm, this is below the minimum

resolution of our fabrication ability and as such the pitch which was selected to be the smallest pitch available.

In addition to resolution related barriers for the fabrication, there is something of a practical limit on the metal thickness since it is challenging to carve deep, narrow trenches using focussed ion beam (FIB). For simplicity it was decided that the metal thickness should be comparable to the grating pitch.

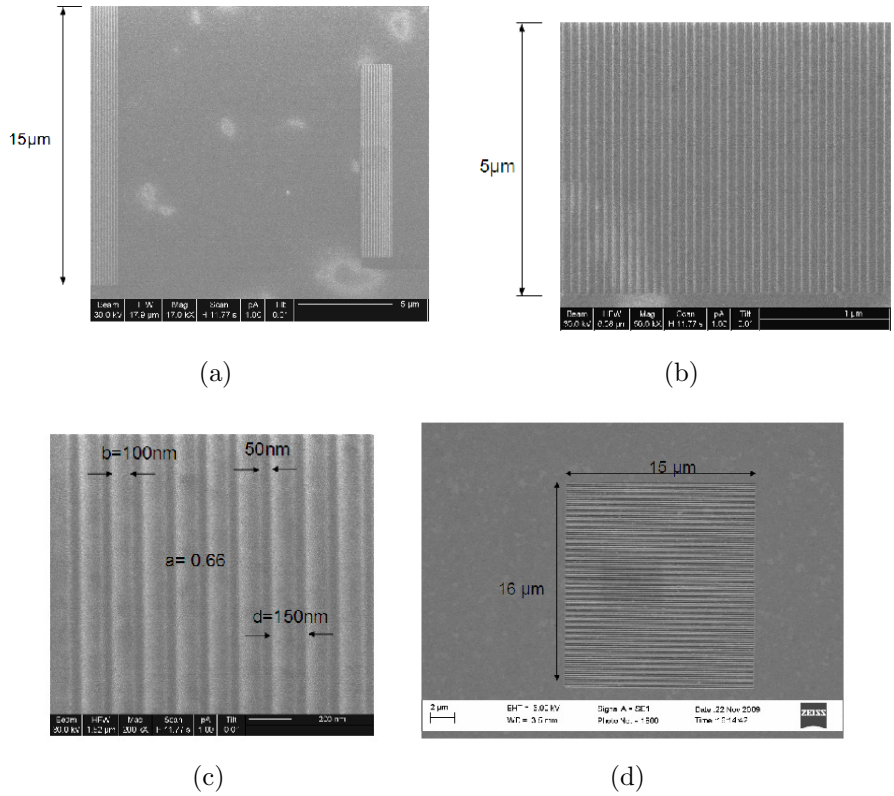
Figure 3.12(b) shows how the fill ratio affects the PER. The fill ratio is relative width of a wire compared to the wire spacing. An optimum value was found at 0.7 which agrees with the optimal region found in [117].



(a) (b)

Figure 3.12: Results of FDTD simulation of the gratings plotting PER (a): against d for varying h with a=0.5 and (b): against a with h=140nm, d=100nm.

Gratings were then manufactured by FIB etching trenches into a 200nm gold film on a glass substrate. An image of an example grating is shown in figure 3.13.



(a) (b)

(c) (d)

Figure 3.13: FIB (a)(b)(c) and SEM (d) Images of a FIB etched polarizer with the parameters shown on (c)

3.2.1 Polarization Measurement Apparatus

An apparatus was constructed to measure the PER of the gratings, this is displayed in figure 3.14. The principle is that a laser beam is polarized and focused down to a diffraction limited spot on the grating, the transmitted light is then collected into a power meter. Rotation of a $\frac{\lambda}{2}$ waveplate (HWP) before the sample rotates the incident polarization and the relationship between the polarization angle and the transmittance of the grating is measured. Many gratings were etched onto a sample and these were found by illuminating with a diffuse white light via a flip mirror (FM) and using a 92:8 beam splitter (to reduce the power loss of the transmitted laser beam) imaged onto a camera. Sample images from the apparatus can be seen in figure 3.16.

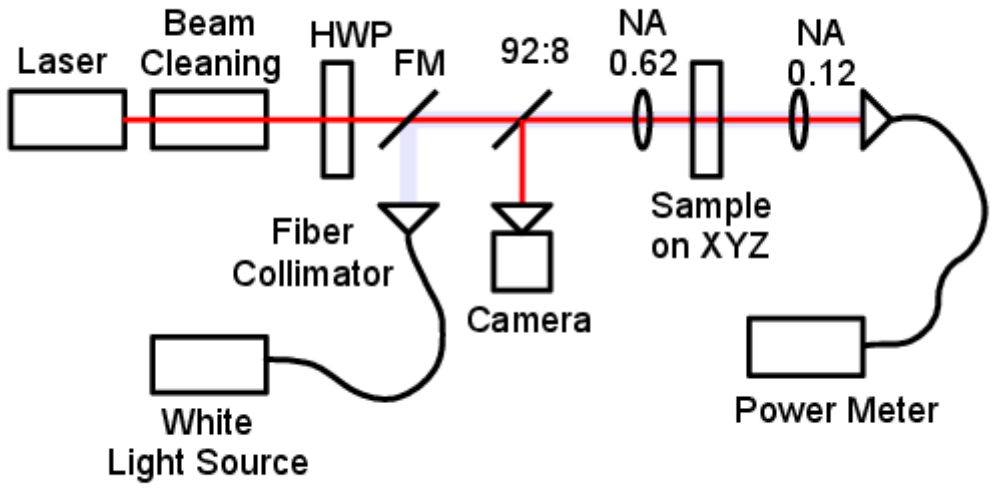


Figure 3.14: Schematic of the apparatus used to measure the PER of gratings. Abbreviations used are HWP: Half Wave Plate and FM: Flip Mirror. Beam cleaning consists of an anamorphic prism pair to render the elliptical beam into a circular beam and then a beam expander with a pinhole to spatially filter the beam to a Gaussian and expand it to fill the back aperture of the focusing lens.

In order to ensure that as little stray light was measured as possible, it was arranged that the laser spot size was smaller than the size of the grating by the Rayleigh Criterion [118]:

$$\sin\theta = \frac{1.22\lambda}{D} \tag{3.5}$$

$$d_{spot} = \frac{1.22f\lambda}{D} \tag{3.6}$$

$$= \frac{1.22\lambda}{2NA} \tag{3.7}$$

the NA 0.62 lens yields a diffraction limit of 622.8nm for 633nm light.

A set of gratings of $d = h = 100nm$ with varying values of a was fabricated and the results are shown in figure 3.15.

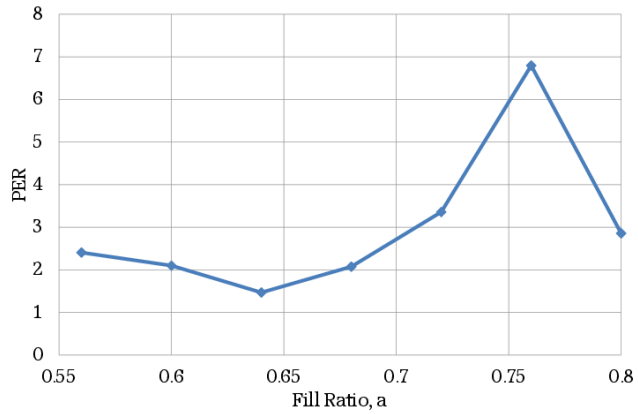
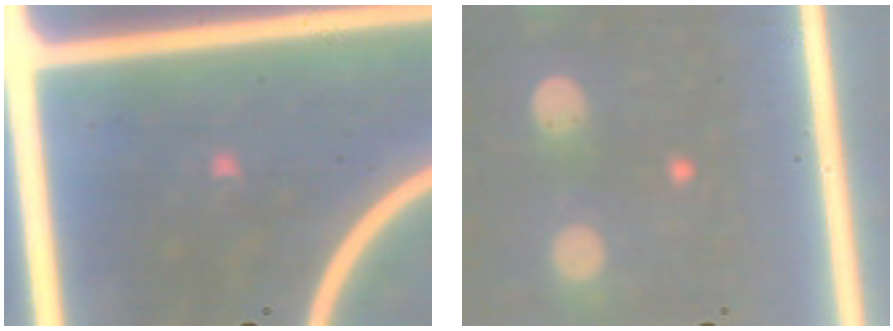


Figure 3.15: Measured PER with varying values of metal fill ratio

Significant difficulty was found etching narrow, deep gratings which degraded the measured polarizing effect over a wide range of grating parameters. The shape of the graph in figure 3.15 is similar to the simulation from figure 3.12(b) albeit with the reduced peak of just below 7:1 moved from $a = 0.7$ to $a = 0.76$. It is thought this degradation is due to light leaking through the gold film which was thinned during the etching process and the grating not having an exactly square profile.



(a) (b)

Figure 3.16: Sample Images from the Polarizer/Plasmon measuring apparatus. (a) shows $5\mu m$ wide alignment structures used for finding the structures. (b) shows two nanoholes. The central, red spot is the focus of the laser which was aligned such that it was fixed in the centre of the image.

3.2.2 Polarizers Summary

The process of etching nanoscale wire grid polarizers shows promise and should be pursued further from these proof-of-principle experiments. Clearly however in its current state this technology is not sufficiently mature for integration into the QKD system of this thesis.

One possible improvement would be to change the metal used for the wires to something with a higher absorption coefficient at the wavelength in question (currently 633nm). The absorption coefficient, χ [119]:

$$\chi = \frac{4\pi\kappa}{\lambda} \tag{3.8}$$

where λ is the photon wavelength and κ is the extinction coefficient, which is the imaginary part of the refractive index of the material. χ is also defined such that $1/\chi$ is the distance into the material where the energy density has fallen to $1/e$ of its initial value [119].

Some values of χ for various different metals at 2eV (619.9nm)[2] are given in table 3.2 which show that the extinction coefficient for Aluminium is higher than that of Gold which was used in this investigation and Silver which was unsuccessfully attempted.

Metal	κ(@2eV)	$\chi\,(10^7 m^{-1})$
Ag	4.18	8.47
Al	7.479	15.16
Au	3.16	6.41
Cr	4.36	8.84
Cu	3.24	6.57
Ni	3.65	7.40
Pt	4.07	8.25

Table 3.2: Extinction and Absorption Coefficients for various metals at 619.9nm. Values from [120]

3.2.3 Plasmonic Gratings

In addition to the FIB etched polarizers, an investigation into annular gratings was performed using a modification on the apparatus in figure 3.14 (modified diagram in figure 3.17). The principle behind the annular gratings is that they can couple light efficiently into surface plasmons (electronic excitations) on a metal substrate and then, providing the dimensions are chosen correctly, interfere constructively and couple back into photons in a central hole. The extraordinary transmission via plasmon excitation was first observed in hole arrays [121] and slits [122] however it is also possible in a circularly symmetrical system [123, 124] which is more useful for point source applications.

The measurement apparatus was a modified version of that depicted in figure 3.14 with the following modifications: The white light source fiber was replaced with 4.3μm core single mode fiber (SMF); the fiber collimating lens changed from

[2]closest available value to 633nm

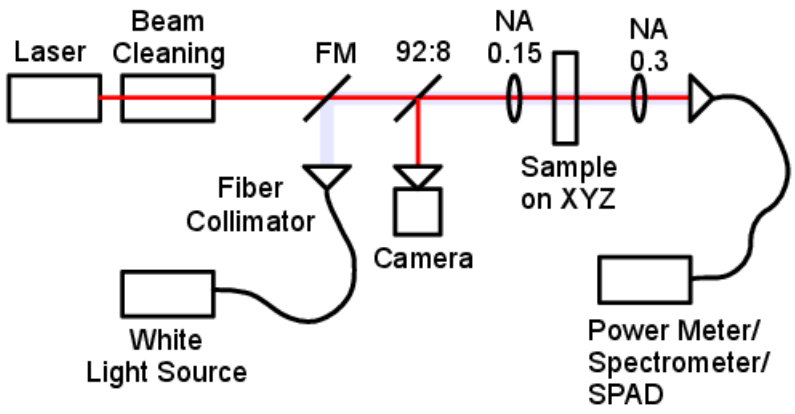


Figure 3.17: Schematic of the apparatus used to measure the spectral dependence of the optical transmission of Plasmonic gratings. Details described in figure 3.14, modifications to that figure discussed in the text.

11mm focal length, NA 0.25 to 4.43mm focal length, NA 0.56; the collection lens changed from 5x microscope objective to 10x microscope objective; the redundant $\frac{\lambda}{2}$ plate removed; single photon avalanche diode (SPAD) module (for low light power measurements) and Spectrometer used for measurements.

These modifications were made so that the white light spot size on the sample was of the order of the size of the nanogratings, this was calculated by determining the Magnification of the fiber collimator/focusing lens system [87] by:

$$M = \frac{f_f}{f_c} d_{spot} = M d_{fiber} \tag{3.9}$$

giving $d_{spot} = 17.86\mu m$ of single mode light. Single mode light is necessary since plasmon interference is required at the nanohole so coherent excitation is required.

The expected result was to see an enhancement of the transmission of light at a wavelength of half the grating period, this was not observed possibly due to damage while imaging using the single beam FIB or due to the low resolution of the

spectrometer. Further information on this study can be found in the Masters thesis of Lifeng Chen [125], the student for whom the apparatus was modified.

3.3 Summary

This chapter has introduced two ideas which will make a miniaturized Alice QKD device more viable.

- Micro polarizers to allow for much smaller LED sizes.
- Simpler beam combining requiring one less degree of freedom in alignment.

In addition to this an apparatus was devised to analyse polarizing effects of microstructures. This apparatus was also modified to analyse extraordinary transmission through small holes surrounded by annular gratings due to plasmonic contributions. The modifications were carried out by the author however the investigation was not hence it has not been covered in this thesis.

The work was carried out into developing micro polarizers has proved promising however the devices are not yet useful for integration into the QKD system as a whole.

Chapter 4

The Bob Terminal

The Bob terminal comprises apparatus for the discrimination and detection of the polarized photons transmitted from the Alice device, the time tagging of these detections and the offline processing of the tag data to synchronise and generate key. The terminal also contains the means for an unsecured public channel to be established between Alice and Bob (wifi or ethernet). These components are described in more detail in section 2.1.2 and section 2.2.2.

Externally the terminal has a protruding shelf with a magnetic kinematic mount for repeatable realignment of the Alice device to the Bob optics. Initially this process was going to take place with a hand-held Alice device utilizing a second optical path along which a (spatial) alignment pattern was displayed, key would only be generated when the alignment pattern fell within a designated sensor area on the Bob device. For now the magnetic dock is sufficient to display the possibility of an instantly dockable device without the added complexity of the active alignment.

Figure 4.1 shows the current state of the Bob Terminal. This box contains all of the blue elements from the image at the beginning of chapter 2 (figure 2.1). Also visible is the shelf on the front containing the magnetic mount used for docking the Alice device to the system. This device has not changed superficially since chapter

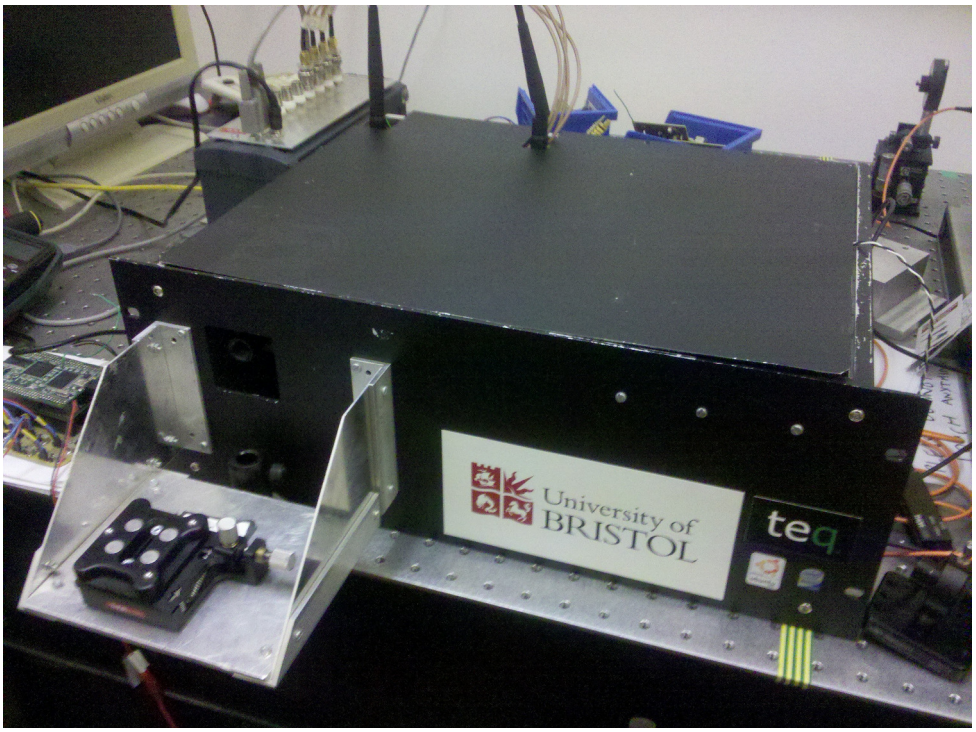


Figure 4.1: The Bob Terminal

2 and will probably not change until the Alice device requires a change in mounting method. That said, internally there are many improvements that can be carried out to improve the performance of the system, this chapter will deal with some of them.

4.1 Requirements

For the system described in chapter 2 to be commercially viable, the speed of key generation needs to be increased and the system must be checked for conformity in light of new security proofs and attacks such as [48, 67]. Speed has been partially and indirectly addressed earlier in that the smaller LED arrays mentioned in chapter 3 will have a lower capacitance [126] and therefore will be able to be driven at higher rates and with shorter pulses [127], as a partner to the transmission speed increase, an upgrading of the ability of the detectors to detect fast signals accurately and efficiently is necessary.

4.2 Ideas

4.2.1 Active Quench

With the dead time of the detectors as it is ($\sim 3\mu$s), operating at 10MHz, a significant proportion of photons are lost while the detector is recharging (60 clocks/dead time). In fact, the situation is even worse when one considers "blinding"; whether malicious or not, due to the assumptions made in section 1.5; that is, that if a detector detects a signal while another detector is recharging, it must be discarded since there was a bias present in the detection probabilities which could be a manifestation of an eavesdropping attempt. [41]

A simple countermeasure is merely to make the recharge time very small such that there is a low probability that there will be another detection in the period when there is bias in the detectors. Whilst one might naïvely think that since dead time τ_D decreases for decreasing values of load resistor R_L then one could pick an arbitrarily small R_L and the dead time would decrease consequently. However, since the current through the single photon avalanche diode (SPAD) must be below a threshold current I_{LATCH}, specific to the device, for spontaneous quenching to occur, a resistor which does not limit the avalanche current sufficiently will either lead to an increased, erratic dead time as statistical fluctuations are required to quench the avalanche or, worse, a reduction in current due to the increased resistance of the SPAD heating, a situation that could lead to catastrophic failure of the device [128].

What is necessary then, is some sort of active solution to quenching which detects the onset of an avalanche and reacts with some pulse to quench the avalanche and allow rapid recharging through a small R_L [97]. As can be seen by comparing the passive circuit depicted in figure 2.8 to the proposed circuit in figure 4.2 that it can be a relatively simple modification.

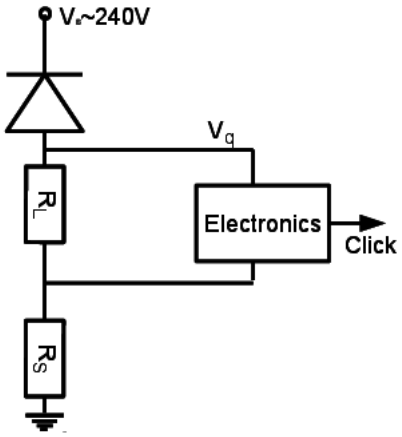


Figure 4.2: Simplified diagram of Active Quenching a SPAD, described in section 4.2.1.

The manner by which the active circuit in figure 4.2 quenches the pulse is simply by allowing the SPAD to breakdown sufficiently to trigger a comparator and then using the comparator output to trigger a Quenching Voltage V_Q such that $V_Q > V_R - V_B$ (where V_R and V_B, which were discussed in section 2.2.2.2 are the reverse bias voltage and avalanche photodiode (APD) reverse breakdown voltage, respectively) which stops the avalanche by reducing the voltage across the SPAD to lower than its breakdown voltage. After a given time, to allow all carriers in the junction to dissipate, the V_Q is removed and the SPAD charges through R_L and is ready to detect again.

4.2.2 Active Recharge

Noting that recharge time is dependent on the RC of the charging circuit made up from the resistive load that causes the quench and the capacitive contributions from the SPAD junction capacitance and the quenching circuit stray capacitance (both of the order of pF) a further modification can be performed which, upon the removal of the quenching voltage, bypasses R_L and connects the low side of the SPAD directly

to ground resulting in a very rapid charge since the only resistive contribution to RC then comes from the internal resistance of the diode. It is important to choose the duration of this connection conservatively since a breakdown during this period could cause undesired behaviour. The combination of active quenching and an active recharge/reset is referred to as an AQAR configuration. This feature is schematically shown in figure 4.3.

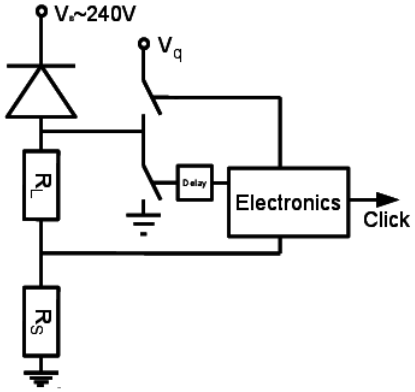


Figure 4.3: An Active Quench Active Reset circuit, described in section 4.2.2.

4.2.3 Active Holdoff

An issue with free-running detection schemes is that after a detection in a basis, the detection efficiency of the two states within the basis are unbalanced (ideally for a perfect AQAR system this goes from 50:50 to 100:0 and back again during one dead time, for passive quenching it is a continuous change). Regardless of the quench/recharge form any detection in this window could leak information to the eavesdropper since she can be assumed to have full access to the public discussion and therefore the timing information. This is even more problematic when one considers that these invalid bits could "pile up" and for any given train of detections where there is less than one whole dead time between any two bits only one sifted bit is extractable, for increasing transmission rate this causes the secret bit rate to tend

towards zero. This issue is discussed in [129] which parameterises the effect and suggests a limitation on the transmission rate as:

$$\rho_{TX}^{max} = \frac{5.92}{8p\tau} \tag{4.1}$$

where p is the probability a sifted bit is detected per transmission rate clock and τ is the detector dead time. And further arguments as to limitations and security are made in [130] A far more simple and effective approach is to enact a hardware implementation, disabling all detectors for τ after a detection is made on any detector [131]. A modification to figure 4.3 for a two detector system is shown in (figure 4.4).

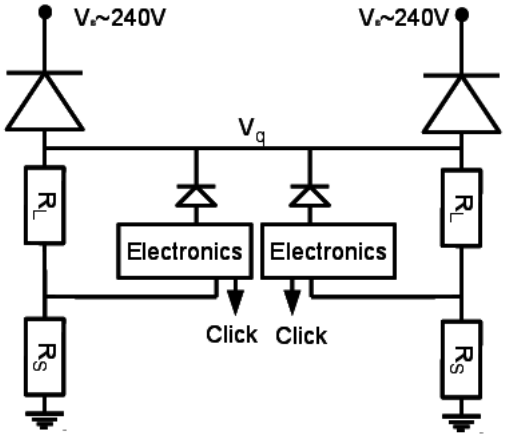


Figure 4.4: Modification to the Active Quenching to allow for an Active Holdoff

A Monte-Carlo simulation was performed wherein two detectors were simulated and were given realistic chances of triggering per clock (μ of 0.3 and Bob efficiency of 30%) and a counter was implemented to monitor dead time. If a click occurred for one detector, it was only recorded if the counter indicated both detectors were active. Otherwise it was discarded and the counter reset. To simulate the Active

Holdoff the behaviour was modified such that when a click occurred while a detector was dead, it was still discarded however the counter was not reset (since the opposite detector would be quenched also).

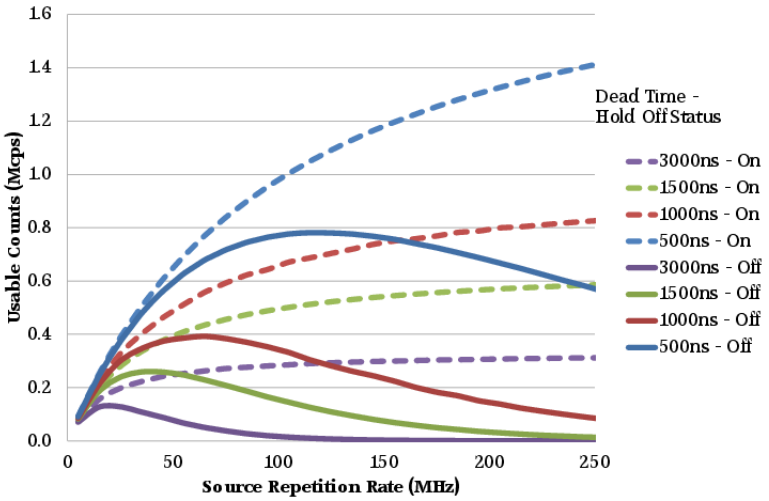


Figure 4.5: Usable count rates of a two detector system for varying dead times when the active holdoff is disengaged (solid lines) and engaged (dotted lines).

As can be seen from figure 4.5, for a non-holdoff system (solid lines), there is a peak repetition rate for a given dead time which corresponds to when the "pile up" of detections overlapping the dead times of previous detections becomes dominant and the rate eventially decreases to zero. The dotted lines show that when the detectors are all disabled when a click occurs in either, the count rate does not peak, rather it tends towards $1/\tau_D$ (the inverse of the detector dead time) as the repetition rate is increased.

Comparing the two data sets in figure 4.5 the benefit of Active Holdoff can be quantified as the percentage increase of counts when Active Holdoff is enacted versus the situation where it is not. Figure 4.6 shows that Active Holdoff at $\tau_D < 1/f_T$ where f_T is the transmission frequency has a fairly marginal benefit but as the dead time increases the Active Holdoff can detect usable counts long after the standard system has become saturated. This improvement is demonstrated in figure 4.7 as the percentage improvement of the count rate for various transmission rates at various

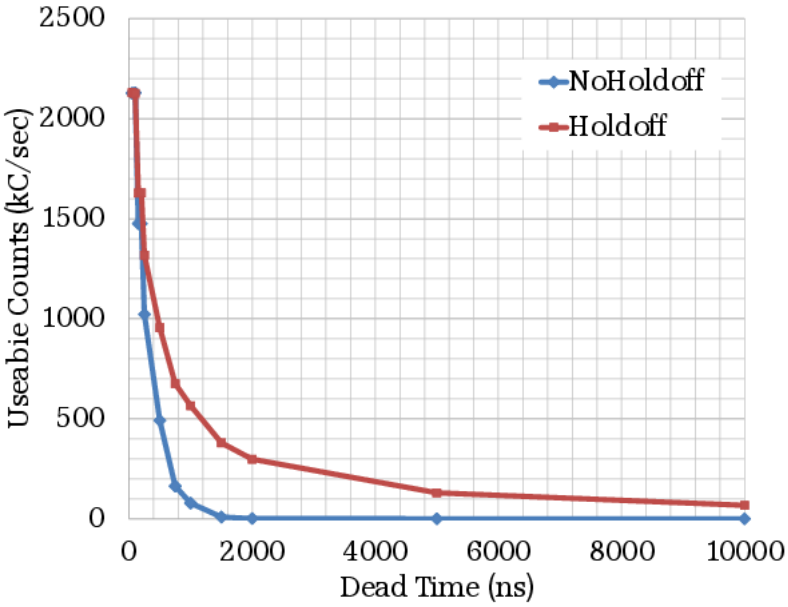


Figure 4.6: Simulation of a system running at 10MHz with varying detector dead times with Active Quenching, to demonstrate the benefit of Active Holdoff

detector dead times, it illustrates the point that the advantage of active holdoff is mainly when the repetition rate of the source is shorter than the dead time and that before this point the active holdoff scheme has no effect.

4.2.4 Cooling

The SPADs are mounted on two stage thermoelectric coolers to reduce the dark count rate [132]. These are single voltage devices and must be driven carefully to maintain a constant temperature since the APD breakdown voltage varies with temperature, which would effect the efficiency and timing characteristics of the detectors. As such a thermistor is present in the package to monitor temperature, an appropriate circuit can monitor this thermistor and modulate the cooling accordingly. In the system depicted in chapter 2 the monitoring was performed by a PIC microcontroller measuring the resistance of the thermistor and switching cooling transistors when the temperature was over a threshold (figure 4.8).

This system was flawed, due to the inability to effectively tune the threshold

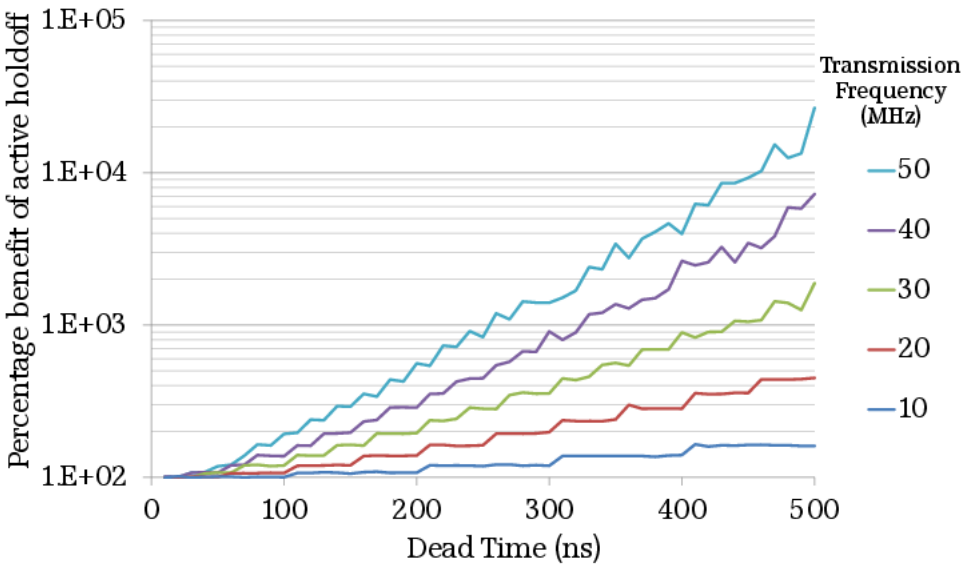


Figure 4.7: Simulation to quantify the increase in usable counts for a variety of system specifications. A lower range of dead times is used here compared to figure 4.6 due to the fact that above this some of the higher transmission frequencies were totally saturating the detectors and leading to excessively high percentage increases (since the non active situation was only detecting one count)

and the requirement for two separate supply voltages (5V for logic and 3V for the coolers), this also added bulk to the system as the grounds were isolated using optoisolators. The PIC software was also undocumented and the author unreachable so a decision was made to implement a far simpler system.

The circuit shown in figure 4.9 places the thermistor into a potential divider with the output voltage connected to a comparator, the comparison voltage is then supplied by a second, fixed potential divider. A variable resistor, RV was placed into the reference potential divider circuit to allow for simple setting of the comparison and therefore threshold temperature. The comparator output was then connected to a MOSFET which, when closed completes the cooling circuit and begins to cool the SPAD. A DCDC converter (LDO03C) was used such that the whole system could be powered from a single 5V supply, this also reduced the current draw from a power supply to below 2A.

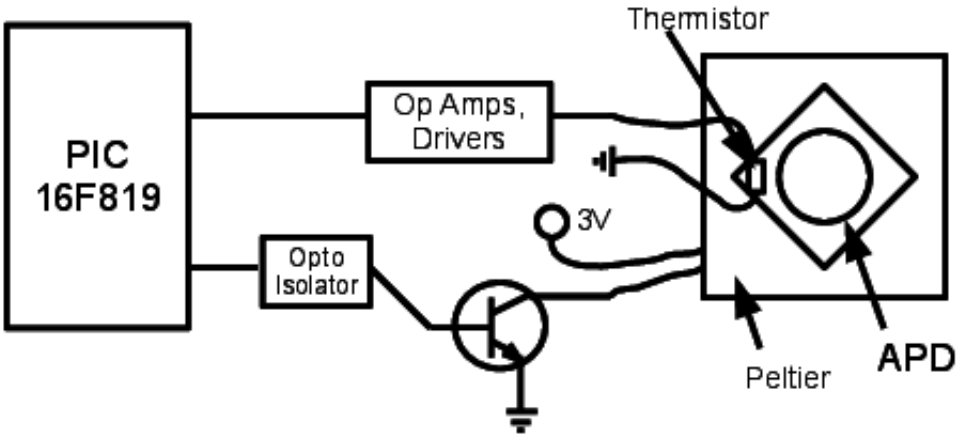


Figure 4.8: The old cooling circuit. A PIC microcontroller was utilized to monitor the thermistor current, switching the peltier coolers appropriately. Optical isolators were used to separate the two supply voltages (5V for logic, 3V for peltier coolers)

The data sheet for the APD[1] gives values for the thermistor as $5.1k\Omega@298K$ and $34.4k\Omega@253K$. The optimal choice of resistor in the potential divider is that which creates the biggest variation in voltage between the quiescent (room temperature) state and the activated (cooled) state

Substituting in R_{Th}, the thermistor values for each state and varying R_1, the fixed resistor the potential divider equation [133]:

$$V_{OUT} = \frac{R_1}{R_{Th} + R_1} V_{IN} \tag{4.2}$$

the widest V_{OUT} swing between 25°C and -25°C is found when R_1 is near to the E24 value of 13kΩ. The graph of this is shown in figure 4.10.

[1]Perkin Elmer C30902S-DTC

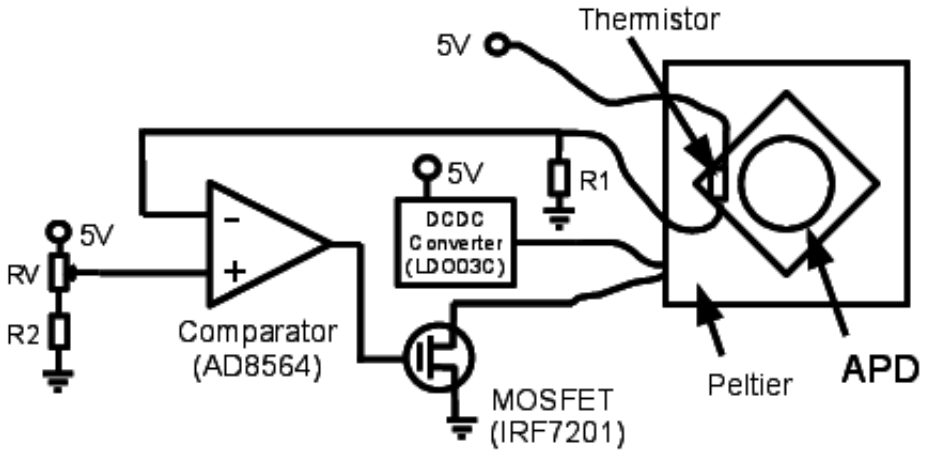


Figure 4.9: The new cooling circuit. The thermistor, R_{Th} is placed in a potential divider circuit with resistor $R1$ which is monitored by a fast comparator (AD8564). The comparator was set such that its output was high when the temperature was above the threshold. The comparator output was used to switch a MOSFET (IRF7201) which supplied power to the peltier cooler. The comparison voltage was set with a second potential divider with $R2 = R1$ thus the value of R_V could be set to be the desired value of R_Th allowing for simple changing of the threshold temperature.

To set the temperature to another value, the thermistor equation must be used [134]:

$$R_{Th}(T) = (R_0)e^{\beta\left(\frac{1}{T}-\frac{1}{T_0}\right)} \tag{4.3}$$

where R_0 is the known resistance at a known temperature, T_0, $R_{Th}(T)$ is the resistance at temperature T and β is a device specific constant given on the data sheet supplied as 3200. A graph of equation 4.3 is shown in figure 4.11.

For small changes in desired temperature, with $R_1 = R_2$, the comparison and thermistor potential dividers are the same and so the comparison trimmer (RV from figure 4.9) can simply be tuned to the value of the desired thermistor resistance

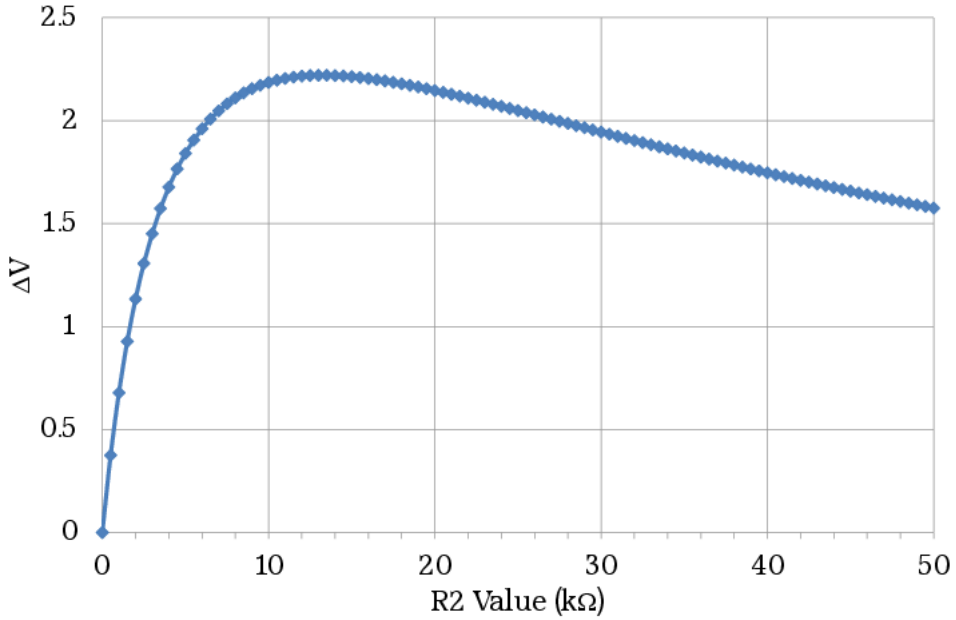


Figure 4.10: Voltage swing at the comparator input between 25°C and −20°C. A maximum can clearly be seen at close to the nearest preferred value of $13k\Omega$

calculated from equation 4.3. For a drastic change in the desired temperature, in order to maintain the widest voltage swing at the input of the comparator, the value of the fixed resistors R_1, R_2 should be recalculated, combining equation 4.2 and equation 4.3 the optimal fixed resistor value in $k\Omega$, for the temperature range between room temperature (298K) and temperature T in Kelvin is given by solving:

$$\frac{d}{dx}5R_1\left(\frac{1}{R_1+5.1}-\frac{1}{R_1+5.1e^{\beta\left(\frac{1}{T}-\frac{1}{298}\right)}}\right)=0 \tag{4.4}$$

which yields the exact value of $R(253K) = 13.253k\Omega$.

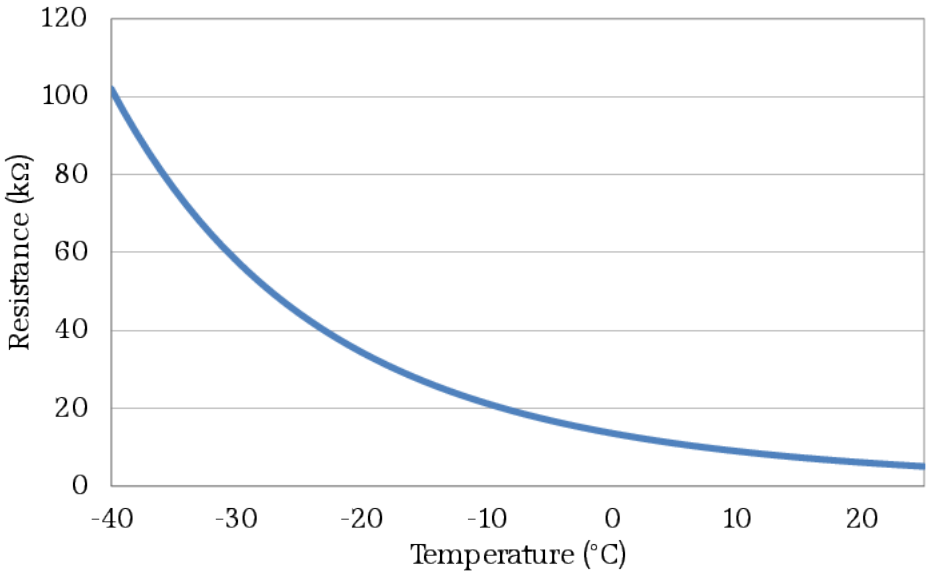


Figure 4.11: Thermistor resistance with temperature for C30902S-DTC APD

4.3 Experiment

4.3.1 Detectors

4.3.1.1 Active Quenching

An active circuit (quench and reset) was designed and produced providing the ability to choose whether either active part is enabled. This allowed for a direct comparison between Passive and Active behaviour since the other circuit characteristics were identical.

There are a range of different active quenching schemes, a large design concern with this application was that the detectors operate at a large overvoltage (25-30V) to provide good timing performance and efficiency [73, 92, 94, 97]

The "voltage mode" (see [97]) detection of the circuit is not the best in terms of timing performance however it was chosen due to the simplicity of the implementation and the similarity to the current passive implementation. This will allow for versatility in testing the QKD performance of the techniques and may help to inform further developments in the quenching scheme to use.

A schematic of the design is shown in figure 4.12. Note that whilst the Active Reset feature is optional, the reset transistor is required anyway to discharge the gate of the quench transistor. It is necessary to employ a diode-OR to isolate the Quench Gate and diode voltage to reset them through the same transistor.

Data was recorded by observing the signal output of the circuit on an oscilloscope and overlaying the pulses following a trigger pulse onto a single trace, figure 4.13. Since dark counts are random, a picture will build up of secondary pulses occurring after the trigger but with no counts within the dead time which can then be measured directly.

Following the dead time measurement, a measurement of the timing jitter was made by employing a fast laser (PicoQuant PDL800B) incident on the SPAD. A

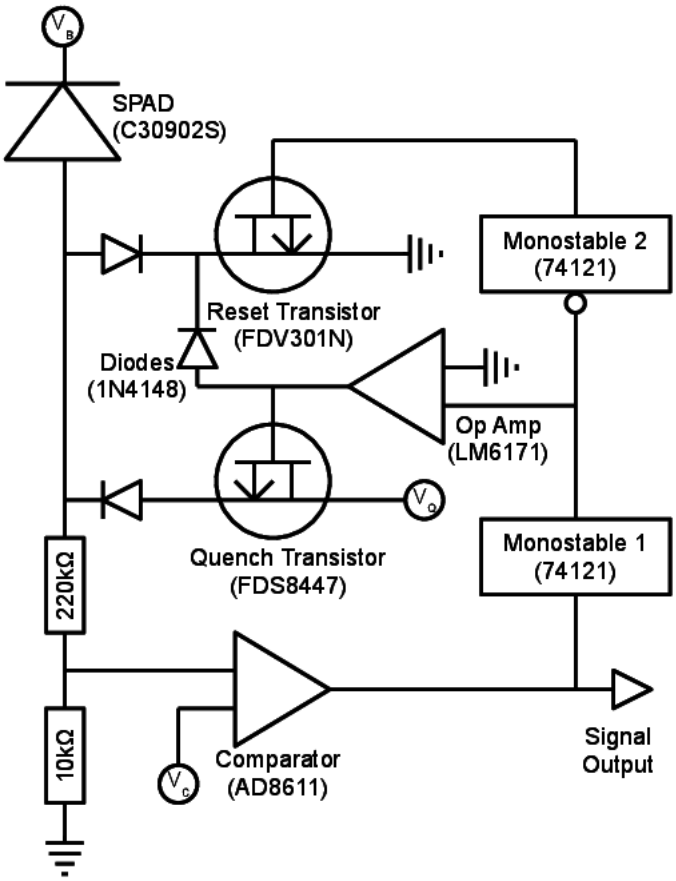
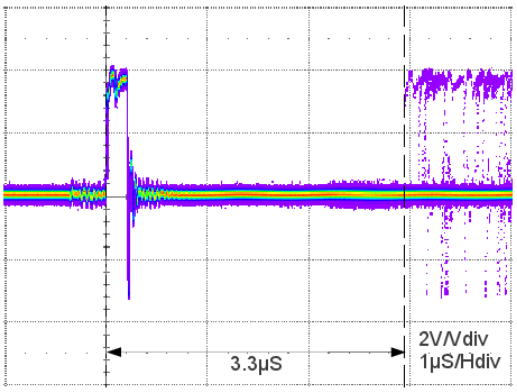
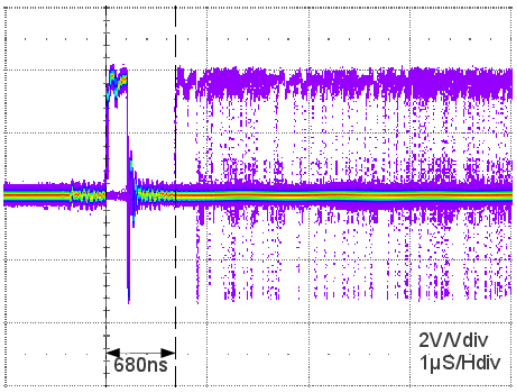


Figure 4.12: Active quenching, active reset circuit. The voltage pulse from a detection is sensed by the comparator which outputs a click synchronous with the avalanche time. This signal also starts a monostable multivibrator which produces a pulse of width independent of the comparator pulse. This monostable pulse is used to trigger a MOSFET which applies a quenching voltage greater than the overvoltage being applied to the APD. A second monostable triggered on the negative edge of the first monostable pulse then shorts the load resistor such that the APD recharges quickly. The gate of the quench MOSFET is also shorted to ensure any remaining charge is dissipated.



(a) Active Off



(b) Active On

Figure 4.13: Oscilloscope traces from which the values of dead time for AQC Circuit with active component disengaged (a) and with active component engaged (b)

.

time correlated single photon counting (TCSPC) measurement was made using a PicoHarp 300 using the laser sync pulse to start the counter and the AQ Circuit output to stop the counter. A histogram of 60 seconds of data with a laser rep rate of 5MHz and a count rate of $1.0 \times 10^5 \pm 0.5$counts/sec from the AQ circuit. figure 4.14

This measurement was performed for a variety of different count rates (by modifying the laser intensity) to investigate the effect of count rate on jitter, this data is displayed in table 4.1. It is suspected that this dependence is a result of the APD

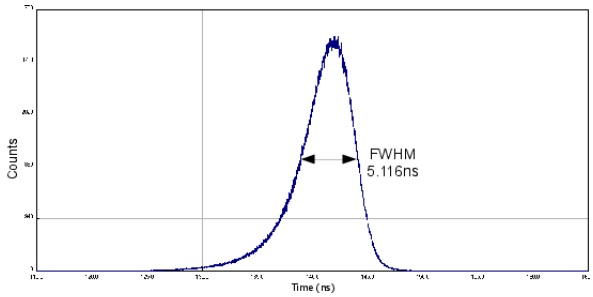


Figure 4.14: Histogram of time differences between Start and Stop of the TCSPC measurement to determine the jitter of the AQ circuit.

bias not recharging fully before the next detection, an effect which is more likely to happen at higher count rates. If the bias has not charged fully then the avalanche peak height will not be as high and a different point on the rising edge slope will be detected. Since the voltage mode rises smoothly (figure 4.15), this can lead to a higher uncertainty in the timing measurement.

Count Rate	Timing Jitter ($\pm$0.140ns)
$5.0 \pm 0.5 \times 10^3$	1.920
$1.0 \pm 0.5 \times 10^4$	1.992
$5.0 \pm 0.5 \times 10^4$	3.076
$1.0 \pm 0.5 \times 10^5$	5.116

Table 4.1: Timing Jitter of AQ Circuit for varying SPAD illumination intensities

The timing jitter is higher than that of a "current mode" passively quenched system however it is sufficient for providing a modest increase to the maximum transmission rate. The absolute maximum timing jitter for a given transmission rate would be such that the timing jitter could lead to a detection being registered as from the preceding or succeeding clock period.

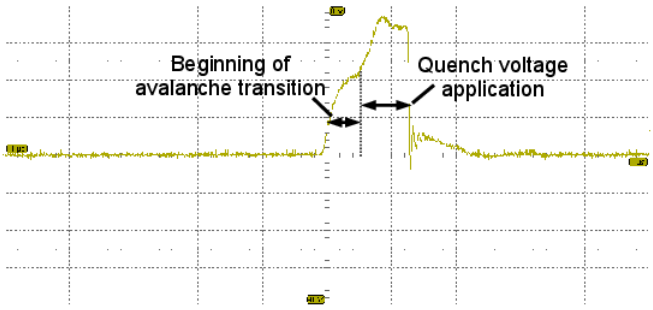


Figure 4.15: The avalanche signal of the active quench circuit. This signal was taken from the anode of the APD rather than the comparator input for clarity; due to the nature of the circuit, the comparator input is an attenuated version of this waveform [97].

4.4 Summary

This chapter has focused on improving the detectors in the system by:

- Investigating Active Quenching of the SPADs.
- Improving the cooling of the SPADs to improve reliability and control.

The SPAD quenching circuit developed is what [97] would refer to as a mixed Passive-Active quench circuit. It achieved a 79% reduction in the dead time (from 3.30 to 0.68 μS). This circuit was not stable enough for inclusion into the system however it is a good starting point for further investigation. A simple active hold-off was attempted by implementing a 4 input OR gate with diodes connecting all quench signals together however this injected spurious signals back into the SPAD terminals which triggered the comparators resulting in the circuit output oscillating by producing repeated quench pulses.

The cooling circuit proved effective and was installed into the final system. The main advantage of this version was that it runs on a single supply voltage (5V) and consistently works[2].

[2]there was some issue with the previous cooling circuit where it would occasionally need to be power cycled to actually commence cooling

Chapter 5

Key Exchange

5.1 Ungated Characterisation

Since the system discussed in chapter 2 was demonstrated, sufficient changes have been applied to the system as a whole that it is necessary to re-characterise the system. There has been a complete overhaul of the field programmable gate array (FPGA) devices in Alice and Bob [79]; the Alice optics were reconstructed in order to attempt to improve a weak LED which was providing difficulty in alignment and the detector cooling was reworked (section 4.2.4) to a simpler method which should prove more reliable. In addition to the changes, the previous characterisation (section 2.3.3) was not as rigorous as it could have been due to time constraints and so a great deal more data was collected for this analysis.

It is unfortunate however that none of the work into the polarizers and LED collimator (chapter 3) or the active quenching of the detectors (chapter 4) was mature or reliable enough to integrate into the quantum key distribution (QKD) system at this time.

5.1.1 Magnetic Docking

The magnetic docking method was analysed previously (section 2.3.3) and the results displayed in figure 2.15 prove the repeatability of placement of the magnetic docking system for the Alice device for a small number of replacements. It does not, however, compare this against the natural, random fluctuation of the signal over time. This was rectified and is shown in figure 5.1. The Alice device was placed in the cradle, aligned and a single LED was set to pulse in order to be able to get extinction via looking at the counts alone. Channel count rates were taken every 10 seconds for 500 seconds and the extinction ratio (ratio of counts in the intended detector in the transmission basis vs ratio in the opposite detector in the basis) calculated, then, for a further 500 seconds, in between every measurement the Alice device was removed and replaced in the magnetic dock.

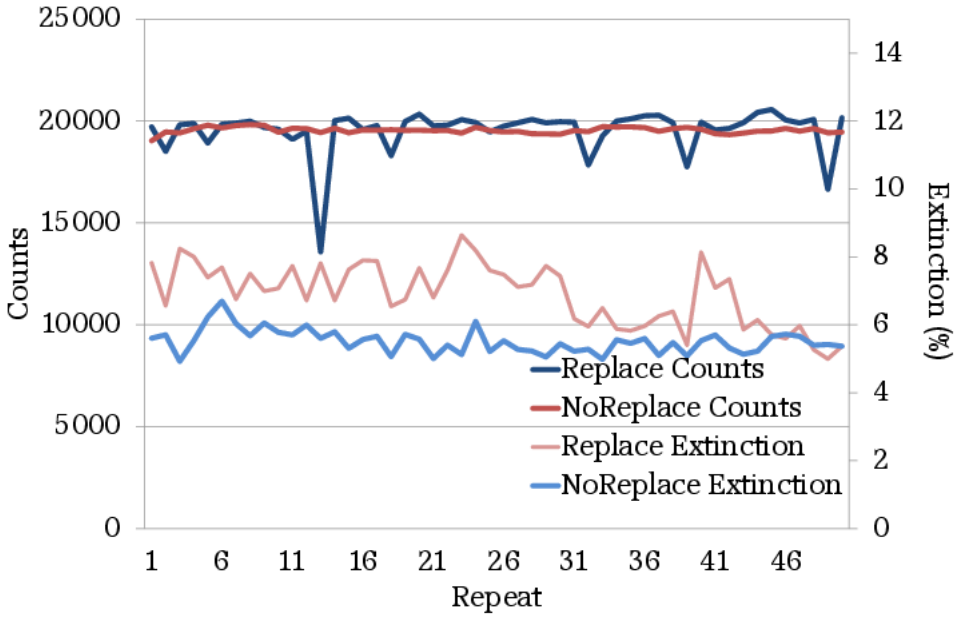


Figure 5.1: The signal count rate and QBER in one basis for subsequent placements of the magnetic mount compared with data taken at the same intervals with the Alice device fixed. The one anomaly in placement had little effect on the extinction and did not effect subsequent placements.

Aside from a notable anomalous point, the replacement count rate remains within ±15% of the mean and 74% of points are within 2.6% of the mean (the maximum deviation for the data where the Alice device was not repeatedly docked). It is important to note that the anomalous point in the replacement data does not correspond to an increase in the extinction ratio and the count rate returns to normal after this, the most likely explanation for this point is that the mount was not correctly docked since it had no effect on subsequent data points. The magnitude of the extinction in this experiment is not important to the characterisation since the experiment was performed with the lights on to allow for easy docking. Characterisation of extinction ratios for each channel separately are found in section 5.1.2 for ungated data and in section 5.2.2 for gated data.

5.1.2 Extinction Matrices

A method of clearly showing the maximum achievable ratios of signal to noise in all 4 channels in one chart is to plot a 4x4 bar chart showing count rates on all 4 detectors for each LED individually. The perfect case here would be 50% counts occurring in the intended channel, 0% in the orthogonal and the remaining split 25% each in the two off-basis detectors.

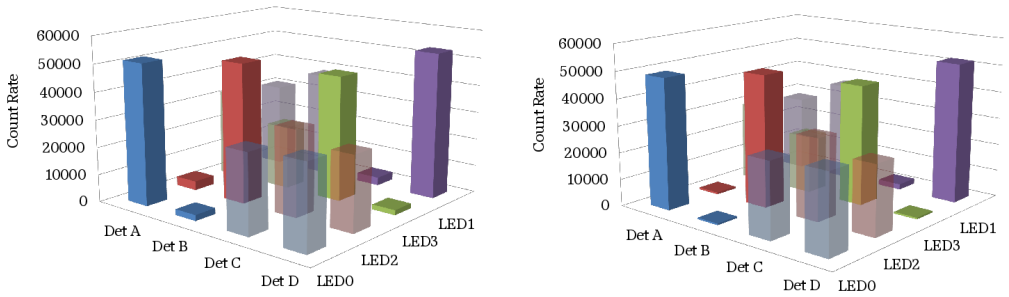
This was investigated by pulsing each LED individually and collecting count rates in all 4 detectors corresponding to each LED. This data is shown in figure 5.2.

Obviously background light, dark counts and imperfections in the polarization (preparation and measurement) will cause some deviation from this situation. A figure of merit here is the “extinction ratio”, the proportion of the counts in the correct basis which was detected in the wrong channel. In the case for LED0, from figure 5.2, it would be:

$$E_0 = \frac{C_B}{C_A + C_B} \tag{5.1}$$

where E_0 is the extinction ratio of the polarization from LED0 and C_A and C_B is the count rate in the detectors A and B.

In order to isolate the optical imperfections (as opposed to the noisy contributions), the count rates when no transmission was occurring were subtracted from all the subsequent count rates and extinction ratios calculated again. The limitations of this method are discussed at the end of section 5.2.2, although it is still useful to establish the contributions to non-optimal extinction from the optics alone.



(a) Extinction matrix

(b) Extinction matrix with dark counts subtracted

Figure 5.2: Ungated extinction matrices of the system with passively quenched detectors. Counts C_{A-D} are shown for detectors Det A-D for each LED0-3 illuminated individually.

LED-Detector Pair	Extinction Ratio	Dark Count Removed
0A	4.17%	1.20%
2B	6.54%	1.43%
3C	4.03%	1.37%
1D	5.54%	3.90%
Average	5.07%	1.98%

Table 5.1: Numerical extinction ratios with and without the channel dark counts.

5.2 Gating and Reconciliation

5.2.1 Transmission Isolation

One of the points in the design brief (table 1.1) is that the establishing of a key should be simple, to this end there is no information regarding the start and stop times exchanged before the quantum transmission. To achieve this in postprocessing, the Alice device instructs the Bob device to start collecting data and then starts transmitting and instructs Bob to stop collecting after the transmission has finished. Thus there is a range of time tags either side of the transmission window where there was no signal to detect. The first step of data analysis is to find the start and stop time tags of the transmission inside Bob's data.

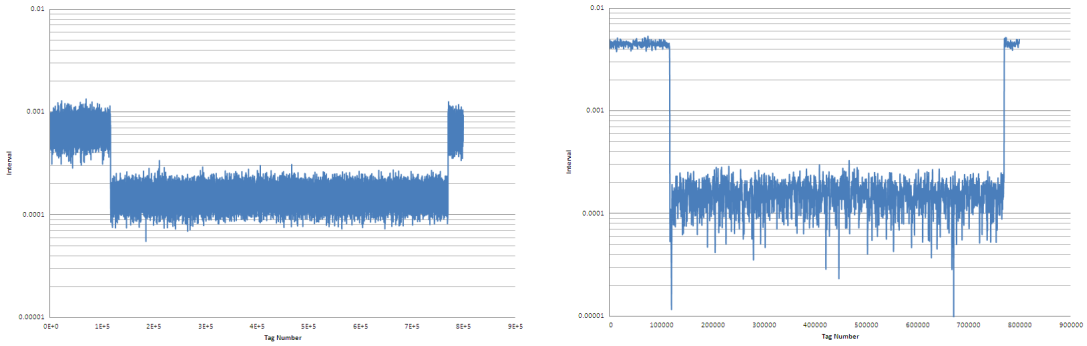
The transmission boundary finding process is carried out by analysing the intervals between successive time tags, the count rate will obviously increase during the transmission and as such the gaps between time tags will decrease. Therefore one can coarsely determine the points at which the transmission starts and stops by looking for sharp changes in a plot of interval against tag number.

This method of interval analysis works acceptably and to within a few tens of microseconds however the synchronisation step detailed later in section 5.2.3 is one of the more time consuming parts of the whole reconciliation process and will, on average, take longer the more imprecise the coarse start point is determined. As such a more complicated and time consuming method of identifying the coarse start point to a high accuracy can actually contribute to a shorter overall processing time.

To improve start point detection, the boundary finding method is changed slightly, the initial search is done by taking the intervals between every nth tag[1] (figure 5.3(a)), smoothing[2] (figure 5.3(b)) to make the transition clearer and then noting

[1] splitting into blocks is used to maintain the speed of the whole technique
[2] Gaussian weighted moving average is used to preserve the shape of the transition

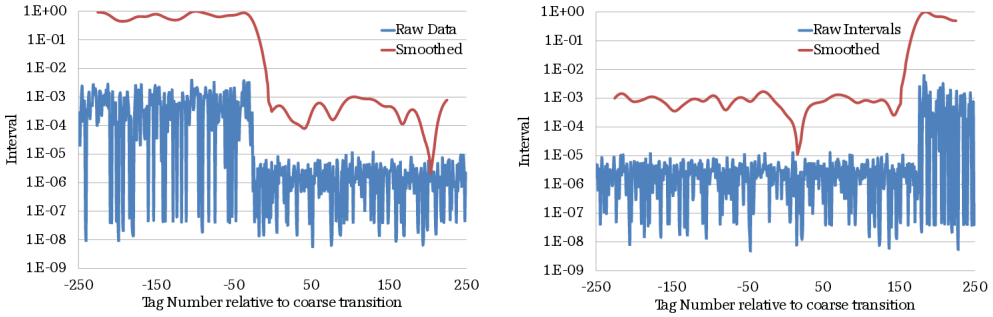
where the threshold is crossed. This finds the start and stop points to an accuracy of approximately $\pm n$ time tags and is referred to as the coarse search.



(a) Coarse Intervals (b) Smoothed Coarse Intervals

Figure 5.3: The coarse boundary finding method, note that in the smoothed case there is a more pronounced transition and less noise "outside" the transmission

Following this step a complete range of tags each side of the coarse start/stop point is isolated from the tags, smoothed as before and a threshold is applied individually[3] (figure 5.4). It is possible the coarse technique identified more than one point at which there was a sharp change in count rate, for now however we are just interested in accurately finding all of the points determined by the coarse method.



(a) Fine Start (b) Fine Stop

Figure 5.4: The start and stop boundaries approximated from figure 5.3 Note the comparison between raw intervals (blue) and smoothed (red)

[3]for especially noisy data, separate thresholds can be applied to start and stop transitions which are lower and higher respectively

As mentioned previously it is possible that the coarse process identified more than one possible transmission window. If this occurred halfway through a transmission it would lead to an incorrect stop being detected and half the data being discarded with a consequent reduction in key length. To remedy this Alice declares her transmission duration and a duration is calculated between each fine start and stop which is compared to the declared transmission duration window, the two starts and stops which give a transmission closest to the declared duration is chosen. An example can be seen in table 5.2 (this example is not connected to the data shown in the rest of this chapter).

Start Tag	Stop Tag	Duration	ΔT
194950	564475	1.80065	2.19935
565200	868750	2.205479	1.794521
194950	868750	4.006129	0.006129

Table 5.2: An example of the Start/Stop finding process which is insensitive to anomalous triggering. "Duration" is the time difference between Start and Stop tags, "ΔT" is the (absolute) difference between that duration and the transmission duration declared by Alice (4 seconds)

5.2.2 Gating

The transmission start and stop times are now determined to within a few microseconds. Using these the transmission can be isolated from the background counts detected before and after the transmission however there is noise present which occurred during the transmission. To reduce the effect of this noise, the arrival times of the tags are analysed and compared to their predicted arrival times, any data outside of the expected arrival time is rejected, this process is referred to as gating and was briefly mentioned in section 2.2.3.2.

The time tags generated by the time interval analyser (TIA) are divided modulo the clock period which can either be found as information shared by Alice or can be determined by FFT of the time tags. The modulus function returns an integer number of clock periods (the tag index) and the remainder (position of tag within clock). The LED pulse width is much less than the clock so a histogram of the remainders (figure 5.5) should present a narrow peak corresponding to the expected arrival time of the signal within the clock period and a smaller background level of randomly distributed noise.

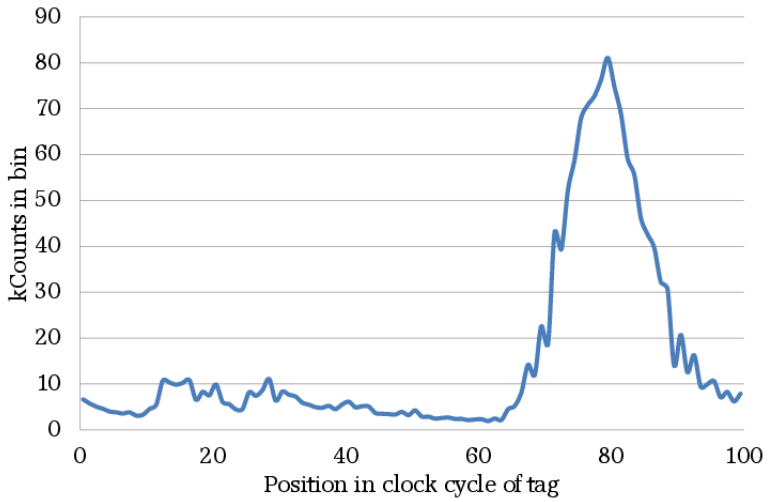


Figure 5.5: Histogram of each time tag in the transmission window divided modulo the clock period (100ns) showing the overall timing jitter of each channel individually and for a stream of random data.

The peak is determined and time tags with a remainder further than some value, the "gate width", from the peak are discarded. The gate width can be arbitrarily chosen as some value close to the LED pulse width (pre measured) or determined from the data by applying successively increasing gates and comparing the proportion of the clock that the gate constitutes against the proportion of the data that the gated subset constitutes. On a plot of P_{Gated} against $P_{GateWidth}$ where:

$$P_{Gated} = \frac{N_{Gated}}{N_{Total}} \tag{5.2}$$

$$P_{GateWidth} = \frac{t_{Gate}}{t_{Clock}} \tag{5.3}$$

where N_{Gated} is the number of the total time tags, N_{Total} which fall within a gate with width t_{Gate} and t_{Clock} is the repetition rate of the source.

The point at which the gradient of this graph passes through unity is the point at which increasing the gate width (and therefore the background, which is uniform across the clock does not increase the signal by a beneficial proportion. This is shown graphically in figure 5.6.

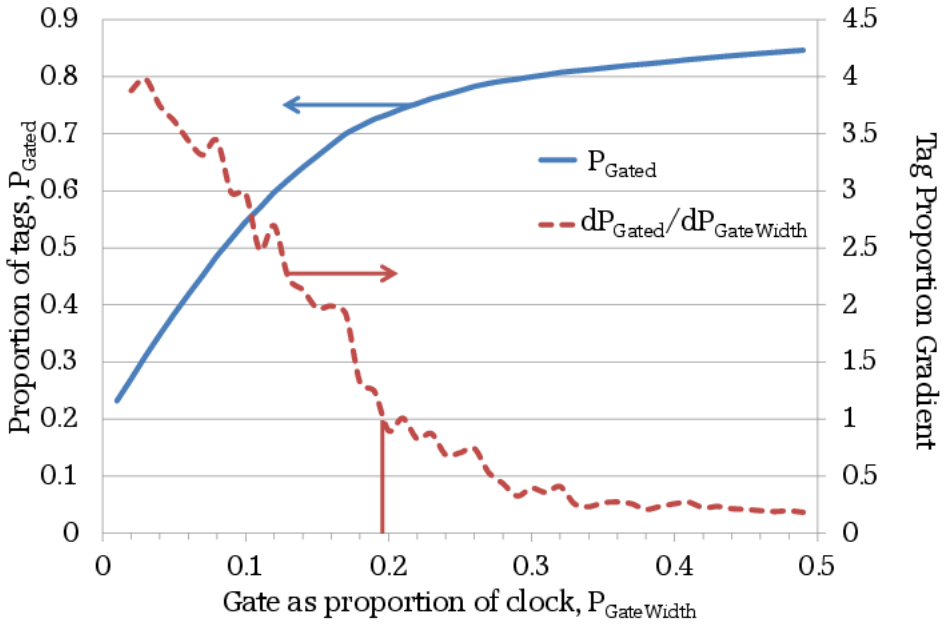


Figure 5.6: The proportion of the signal which falls within a gate of a given width (solid line) and the same data differentiated (dotted line) to allow for easier determination of the point at which widening the gate stops increasing the amount of signal.

Once the data falling outside of the gate is discarded, the remainders of the modulo function can be discarded also and the time tag information can be reduced to a sparse list of integer tag numbers. In a similar manner to figure 5.2, each

LED can then be pulsed separately and the number of gated tags in each channel measured and plotted into an extinction matrix. This data can be seen in figure 5.7 with numerical data in table 5.3.

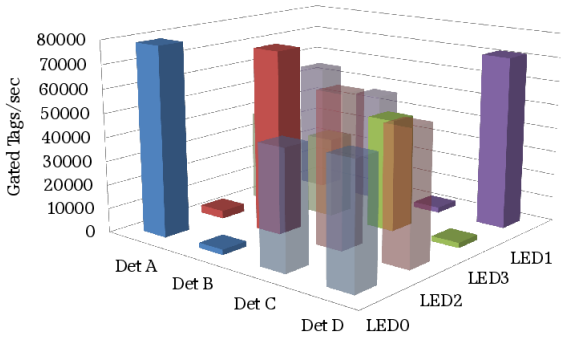


Figure 5.7: Extinction ratios plotted by taking the number of gated time tags (20ns wide gate) from detections in each detector for individually pulsed LEDs.

LED-Detector Pair	Extinction Ratio
0A	2.79%
2B	4.35%
3C	4.67%
1D	3.13%
Average	3.73%

Table 5.3: Extinction ratios of data after gating

Comparing table 5.1 with table 5.3 it can be seen that the extinction ratio of the gated data is better than the ungated when dark counts are included but worse than the ungated without dark counts. This is because, dependent on the gate width, some dark counts will naturally fall into the gate and be counted as signal. The gated data, however, is still a better figure to use since this small contribution of dark counts to the extinction ratio carries through into the actual raw key and it is this value of extinction ratio that should be considered when noting the best attainable QBER.

5.2.2.1 Clock discrepancy

The Alice and Bob devices have the possibility to share a clock over a coaxial cable which is useful for testing however this compromises the design philosophy somewhat and so each device also has an independent oven-controlled crystal oscillator (OCXO)[4] stable to 200ppb. It is however possible that the two clocks are not exactly the same frequency, this will cause Alice and Bob's perception of a time interval to be different and will result in the data appearing to gradually de-sync.

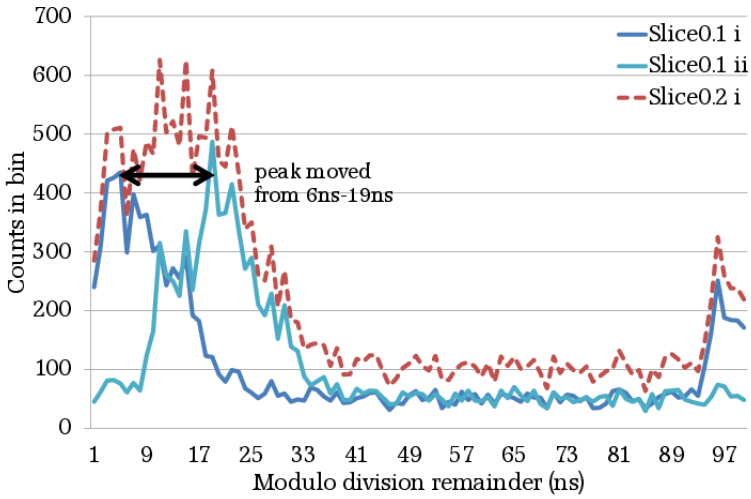


Figure 5.8: A histogram of the time tags divided modulo the clock period for short slices of the whole transmission. Shown here is the peak drift between two subsequent slices (solid blue) and the data from both slices analysed in one go (red dotted), notably the wider slice has a broader peak which would result in requiring a wider gate and thus increase the amount of background light. Numerals in the key refer to the slice number in the transmission.

This will lead to a broadening of the gating peak or in extreme cases an inability to gate the data at all. A solution to this problem is displayed in figure 5.8; the data is split into short "slices" and each is analysed separately. Shown in the figure is a 0.2 second slice (dotted red) and the same slice split in half with the halves analysed separately (solid blue). As one might expect the 0.2 second slice histogram is the sum of the first two 0.1 second slice histograms, thus the peak is broader.

[4]Micro Crystal HCMOS OCXO

For now, the solution is to simply slice the data into sufficiently small slices (100ms) such that each slice can be gated separately and then recombining the slices once gated. A more in depth method which automatically calibrates the tags is described in section 5.3.1.

5.2.3 Synchronisation

It is highly unlikely that the first gated tag corresponds to the first quantum bit (qubit) sent; in order to synchronise the tags to the transmitted data, a small proportion of the sent data[5] is declared (which is then considered public knowledge and cannot be used in the final key), for now, 10%. Each value of the received data is compared to its corresponding value in the list of sent data and the proportion of matches is produced. If the data is unsynchronised 25% of the values will match by chance. All of the indices of the sent data are then offset and the process is performed again.

For the correct offset value the match proportion will jump to 50% and the received data is permanently reindexed with these offset values. A better, but more computationally intensive search analyses only those events where the transmission/measurement bases match and provides a match rate of 50% when the wrong offset is chosen and approaching 100% when the correct offset is used[6]. This method provides an estimate of the QBER of the full transmission by reading off the discrepancy of the peak height from 100%. This technique can be shown in figure 5.9 where the QBER is: $(1 - 0.9737) \times 100 = 2.63\%$. Once the correct offset has been found the data is permanently reindexed with the offset values and this constitutes the raw (sifted) key.

[5]the finite key formula in section 1.6.2 should be analysed to determine the proportion which provides the best bound to the QBER without using up too much of the potential key material.

[6]the most efficient situation here is to perform the basis insensitive search until a candidate for the best offset is found and then to analyse that offset with the basis sensitive search.

It was found that the all of the synchronising offset values were near to 0, with most lying within ±1000 and none lying further than 10000 (corresponding to a transmission finding accuracy of better than 1ms). To this end the offsets were increased outwards from zero[7] and the process was abandoned when the offset exceeded 10000.

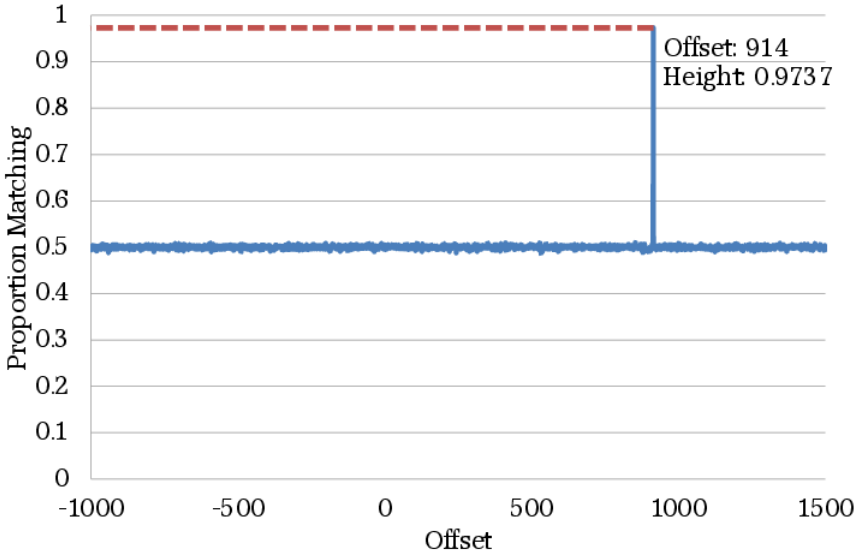


Figure 5.9: The result of searching for the correct offset between the Alice and Bob data using the subset of the data where the transmission and detection bases match. 50% of the data matches by chance however when the synchronising offset is found the match rate increases to nearly 100%. The distance of the peak height from 100% gives an estimate of the QBER. The maximum offset that should be checked before abandoning the search is determined by the efficacy of the transmission finding algorithm (section 5.2.1). It was found that no offset exceeded 10,000 (corresponding to 1ms).

Now that sifted keys can be generated and the errors between them estimated, this is an appropriate time to test the gate width optimization performed in section 5.2.2. Recall that the optimal gate width was determined by finding the point at which widening the gate width by a small proportion produced an increase in tags within the gate of less than that proportion. This method can be checked by performing the synchronisation and error estimation steps repeatedly on the same

[7] i.e. 0, 1, -1, 2, -2

data with different gate widths. As can be seen in figure 5.10 the point at which the rate of increase of tags within a widening gate (dotted line) crosses one is also the point at which there is a sharp change in the QBER.

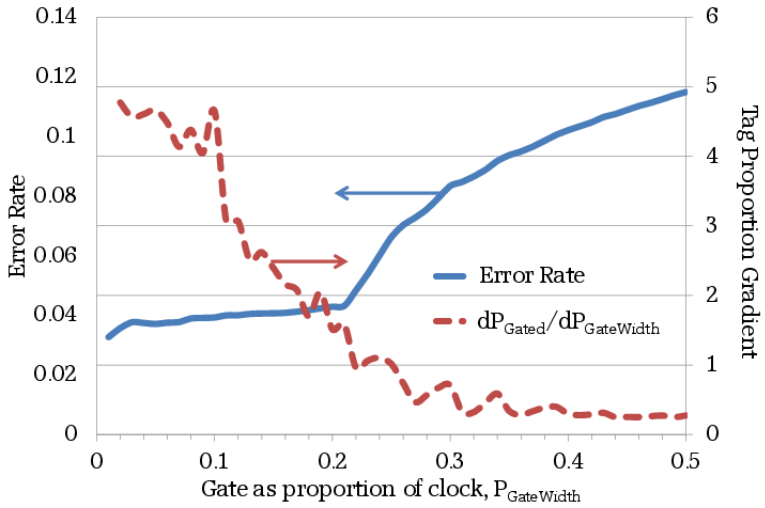


Figure 5.10: The estimated QBER (solid line) as the gate width is changed compared to the method using the proportion of the tags inside the gate (dotted line). The point at which the proportions gradient crosses one signifies when widening the gate by a proportion does not increase the tags inside the gate by a that proportion, note that this point coincides with the sharp increase in QBER

If the gating was required to be performed in slices, it is possible that the clock drift over the whole transmission is longer than the clock period. If this is the case the offset will not be constant for all of the data and each slice should be synchronised separately. This is not as bad as one might think as chances are that the offset will only be ± 1 away from the previous block, assuming this, rather than using 0 as the first guess for the slice offset, the best value for the previous slice can be used. The data for a sliced analysis is displayed in table 5.4. Once the synchronising offset has been found for each slice, each can be applied to the relevant data and then the whole set of slices combined into one synchronised raw key.

Slice	Gate Middle (ns)	Best Offset	QBER (%)
0	4.5	-493	5.85
1	18.5	-493	6.51
2	37.5	-493	7.10
3	41.5	-493	8.34
4	57.5	-493	4.67
5	72.5	-493	7.42
6	80.5	-493	6.33
7	95.5	-493	7.22
8	10.5	-492	5.30
9	27.5	-492	5.58
10	33.5	-492	5.01
11	49.5	-492	6.13
12	64.5	-492	5.80
13	72.5	-492	6.36
14	87.5	-492	5.35
15	4.5	-491	5.45
16	18.5	-491	3.66
17	27.5	-491	5.56
18	41.5	-491	5.64
19	55.5	-491	5.85
20	64.5	-491	5.70
21	80.5	-491	6.88
22	95.5	-491	3.91
23	14.5	-490	7.26
24	18.5	-490	6.65
25	33.5	-490	6.96
26	49.5	-490	5.83
27	57.5	-490	6.58
28	72.5	-490	6.66
29	87.5	-490	5.34
30	95.5	-490	5.33
31	14.5	-489	4.99
32	27.5	-489	5.02
33	41.5	-489	5.79
34	49.5	-489	6.96
35	64.5	-489	5.19
36	81.5	-489	6.48
37	87.5	-489	6.16
38	3.5	-488	6.79
39	18.5	-488	5.42
Average			5.98

Table 5.4: The gate position, sync offset and QBER for each 0.1 second slice of a 4 second transmission. (This is the same data as was used in the example in figure 5.8). Note the gradual increase of the gate middle within the clock period until it reaches the value of the clock period and rolls over, incrementing the data offset value.

5.2.4 Error Correction

Note that currently the raw key still contains errors, values in Alice and Bob's strings which do not match. While this process is generally referred to as Error Correction [135, 136], the process used here merely identifies the positions of these errors and removes them. The challenge in this is to ascertain the mismatched bits without leaking substantial information about the secret bits which do match. The method used here is detailed in [19], the choice of this method here should not be considered a suggestion to its appropriateness in the use model of this project, in fact, the number of separate public communications involved in this method would probably rule it out since the various computational steps would repeatedly find themselves held up by the latency of the public channel.

The process of identifying the errors consists of Alice and Bob both permuting their strings by the same random order, dividing these strings into blocks of length b and comparing the parities of these blocks[8]. If the parities match the blocks are either the same or contain an even number of errors (this issue is addressed later), if the parities do not match, the block is divided in two and the parity of one half is compared. If the parities do not match in this block then the sub block is further divided until the error position is isolated, in the case that the parities of the compared sub block match then the error is necessarily in the other block and the second block can be divided without the parity comparison (which leaks data) needing to be carried out.

A point briefly touched upon in the previous paragraph is that of leaking data, that comparing the parities reveals a small amount of information to an Eavesdropper who must be assumed has some information about the key already from the presence of errors. This is mitigated here by Alice and Bob each discarding one element of any compared block (or sub block). This is where the usefulness of inferring

[8] the sum of all elements modulo 2

a sub block contains errors by virtue of the fact that the other half of the block does not, it reduces the number of comparisons and therefore the number of bits that need to be discarded.

Once the whole sifted key has been compared, Alice and Bob perform a further identical permutation of their data and perform the steps again albeit with a larger block size since the chance of finding two errors in a block of size b has decreased (some errors were removed in the previous step). Randomly permuting the data again should reveal further errors which were previously residing in blocks with an even number of errors. This process of permuting and repeating the algorithm is carried out until the whole string is passed through twice without identifying any errors.

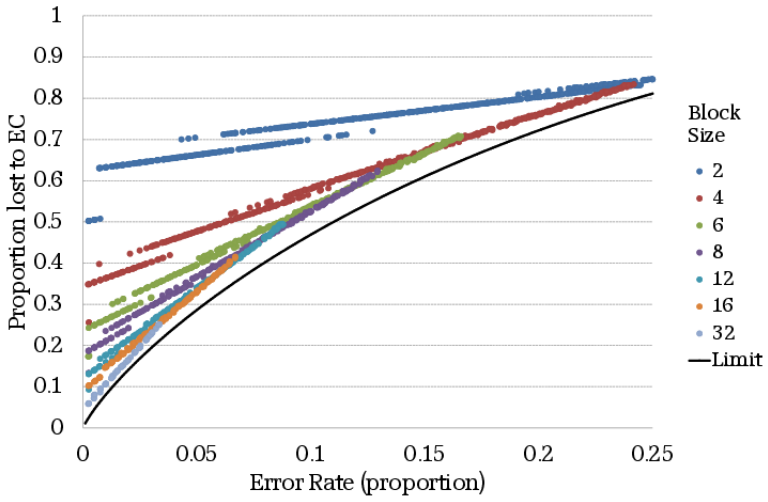


Figure 5.11: The proportion of bits discarded during error correction for a range of block sizes. Each QBER was simulated 5 times, this makes the step-like transition where sometimes the error correction completes in n passes for a given QBER, sometimes $n + 1$. The code efficiency is plotted alongside the Shannon limit, (equation 5.4, the theoretical best performance of an error correcting code.

Figure 5.11 shows the performance of this algorithm for various starting block sizes and compares this to the theoretical lower bound for leakage defined by the Shannon limit at that QBER, $H_2(e)$ where:

$$H_2(x) = -x\log_2(x) - (1-x)\log_2(1-x) \tag{5.4}$$

It can be seen that the algorithm performance tracks the Shannon limit within 20% if the initial block size is chosen correctly.

5.3 Daylight Operation

For further characterisation, data was collected and analysed by the processes above for varying levels of background illumination (provided by a CW red LED powered through a variable resistor) and a plot of QBER against background was obtained (figure 5.12). In the same manner as table 2.1, the system parameters for these experiments are shown in table 5.5.

Repetition Rate	10MHz
Source Wavelength	632.8nm
System Efficiency	10%
Mean Photon Number	0.1
Transmission Time	4 seconds
Gate Width	5ns

Table 5.5: The experimental parameters.

For this process the background was estimated by taking note the number of time tags which did not fall into the gate during the gating process. This is acceptable for laboratory characterisation with a fixed gate width as the relationship between number of ungated tags and brightness (measured by a power meter adjacent to the input lens of the Bob optics) is linear however this is susceptible to attack in a manner similar to the attack described in [49]. The reason the power meter level was not used as an estimate of background is that the power meter has a wider spectral

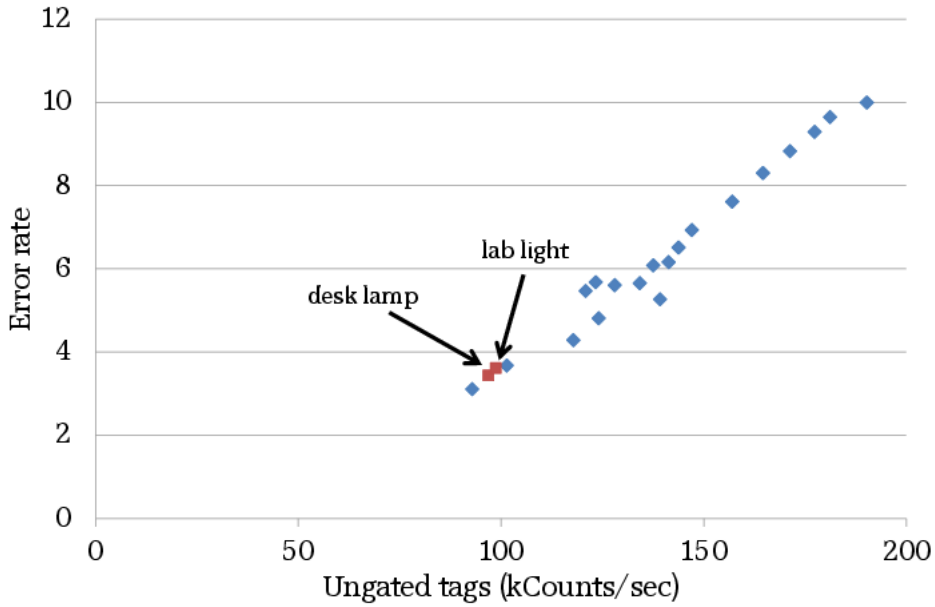


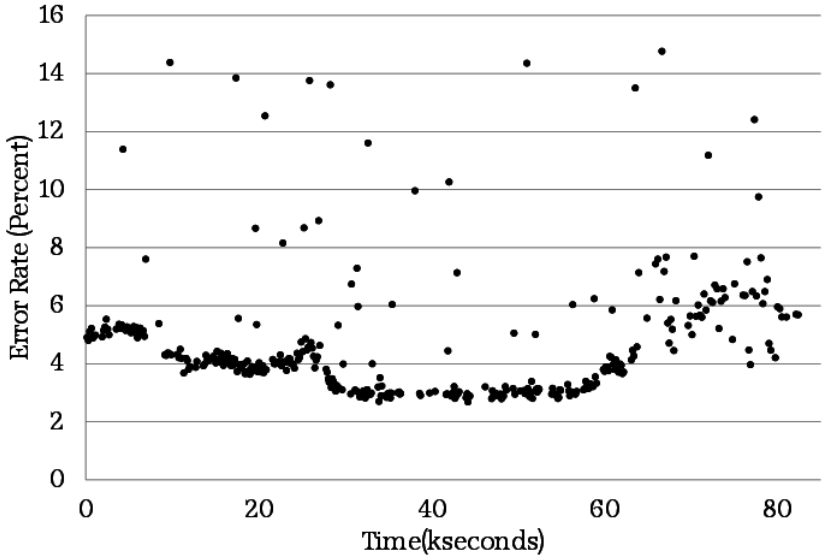
Figure 5.12: The rising QBER with rising background. The blue points signify the data collected by variable LED illumination. The red points are those with more broadband background sources - "desk lamp" being a small desk lamp switched on in an otherwise dark lab, "lab light" being the main lab strip lights switched on.

response than the Bob device which would cause it to overestimate a broadband background compared to a spectrally narrower background. A possible approach to rectify this is to place a photodiode (not a single photon avalanche diode (SPAD)) after the optical filter in the Bob device, taking a small fraction of the light from the optical path with an unbalanced beam splitter for example.

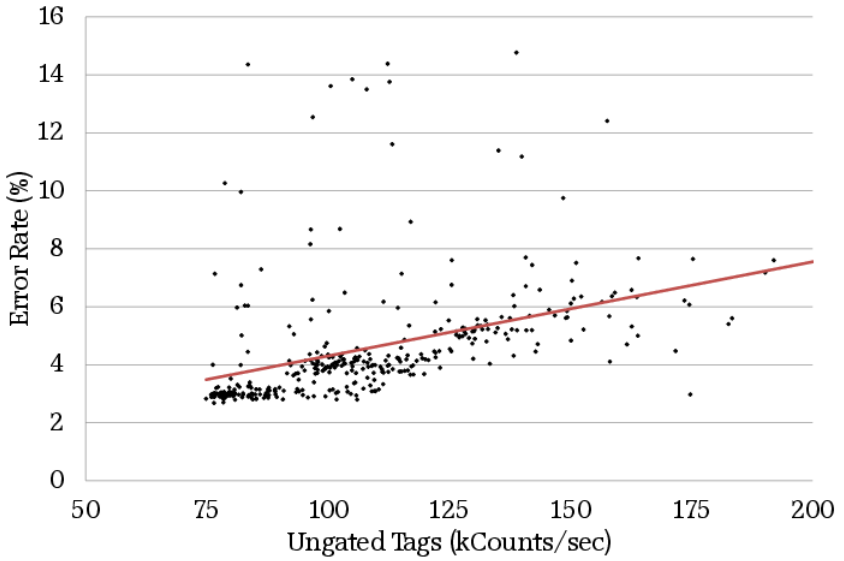
In order to test long-term reliability, the system was also run continuously for 23 hours between 13:00 on 4th June 2013 to 12:00 on 5th June 2013. The system is not intended for constant usage but the process of repeatedly exchanging key shows the repeatability of the system. Figure 5.13(a) shows the estimated QBER for each measurement taken overnight (in seconds since the start of the measurement) and clearly shows a drop in QBER overnight when the background was lower. The sunset time on 04/06/13 was 21:22, the sharp change at 25000sec corresponds to a time of approximately 20:00 which was probably the point at which the sun dropped below

the visible horizon, obscured by trees or buildings.

Figure 5.13(b) plots the same data as figure 5.13(a) but shows the background rate against the QBER, this shows that natural fluctuations in the broadband background light cause the same effect as those in figure 5.12. Note however that the values for background between figures 5.13(b) and 5.12 are not necessarily comparable, this is because the method for estimating the background is not an absolute method and relies on analysing the ungated tags which depends on the gate width.



(a) The QBER at various times through a measurement (in kSeconds) between 13:00 on 04/06/13 and 12:00 on 05/06/13. Note the reduction of QBERcorresponding with night-time. The small proportion of higher error data points can be ascribed to the TIA error mentioned at the end of section 5.3.



(b) A graph of the QBER against the estimated background for each data point from figure 5.13(a). The linear trend conforms to the laboratory testing displayed in figure 5.12

A small proportion ($< 10\%$) of the measurements are erroneously high, this appears to be a fault in the time to digital converter (TDC) on the TIA. This was discovered by noticing the gating histogram (figure 5.5) had developed a secondary peak in the high error measurements. Separating the data into $300ms$ slices and synchronising each individually revealed that there was a gap of several milliseconds in the timing data. Upon checking the Alice electronics and TIA, the TIA was determined to be the source of the error suggesting the timing counter was just not counting over a certain range of numbers properly.

An example of this erroneous behaviour is shown in figure 5.13, each of the slices shown in the figure were synchronised separately (table 5.6) and a drastic jump in the synchronising offset was seen to occur from close to zero to ≈ -40000 clock periods corresponding to a discontinuity of the order milliseconds.

The discovery of this behaviour also highlights the fact that this QKD system is providing a beta testing of the TIA device being designed at Bristol. In the current version of the TIA this fault has been rectified.

Slice	Gate Middle	Best Offset	QBER (%)
1	$10.5ns$	38	1.76
2	$10.5ns$	38	2.51
3	$35.5ns$	-40087	5.36
4	$35.5ns$	-40087	4.34
5	$35.5ns$	-40087	2.05
6	$35.5ns$	-40087	1.30
7	$35.5ns$	-40087	3.35
Average			2.96%

Table 5.6: The first 7 slices of a glitched data set showing how the offset drastically changes and at this point the gate centre moves as the discontinuity is not an integer number of clock periods. The QBER for the non-sliced analysis (using the first 10% of the transmission to sync) for this data set was 11%.

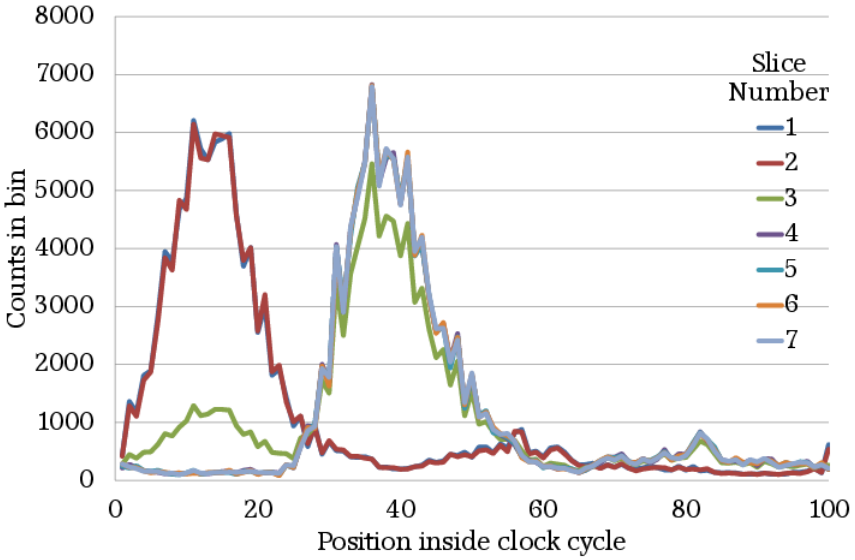


Figure 5.13: A histogram of the remainders of the time tags divided modulo the clock period of the first 2.1 seconds of a transmission separated into 300ms slices. The third slice shows two peaks as the discontinuity occurred within this period, the peak moves as the discontinuity is not necessarily an integer number of clock periods.

5.3.1 New TIA

After finding the fault in the TIA, the opportunity was taken to upgrade to the newest working version of the Bristol-developed TIA (Figure 5.14) which has independent timing circuitry on each input. This allows for more accurate measurements however it introduces an issue in that there may be delays in the different channels which could move the relative positions of the gating histogram for each channel. To this end the time tags were analysed individually to find a range of calibration factors which were then applied to the tags before the analysis described in section 5.2.

Figure 5.14: The new TIA device developed at Bristol. The timing accuracy is improved by about 25% over the previous device. It also now interfaces over USB rather than ethernet and timing bin width calibration is performed onboard.

First the time tags are split according to their channel of arrival, some small time window of tags are taken[9] and divided modulo the clock period (figure 5.15). The data in figure 5.15 shows that channels A&C and channels B&D are synchronised to each other but there is some difference between these pairs.

Due to the unsycnchronised clocks, the histograms are quite noisy and therefore some way of determining the middle of the peaks is required. Fitting a gaussian to the data appears to be a promising idea however if the peak were close to 0 or

[9]in order to reduce the effects of the unsynchronised clocks

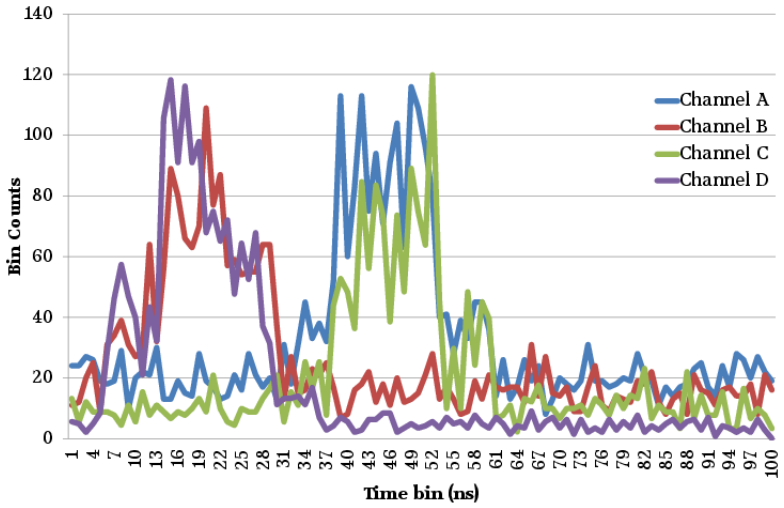


Figure 5.15: Histogram of tags divided modulo clock period from the first 100ms slice of a transmission. Differing delays in the pulse discrimination circuit and the TDC cause the peaks to appear in different places. This confuses gating and is detrimental to QKD performance.

the clock period then the fitting algorithm may produce odd results due to the data "wrapping around" to the other side of the graph. Thus, in order to fit a gaussian, a rough estimate of the middle of the peak needs to be made so the data can be offset such that it is close to the middle. This is performed by assuming that the largest bin in the data set roughly corresponds to the middle and adding this value to all of the time tags in the channel and will be known as the rough correction. This process is shown in figure 5.16 and the gaussian fits are shown in figure 5.17. The gaussian fits now are also not exactly in the middle however the peak fitting algorithm returns the value of the peak middle, this value is known as the fine correction.

Adding the fine correction to every tag will then bring all of the channels into synchronisation which can be verified by applying the corrections to some other temporal slice of tags in the transmission figure 5.18.

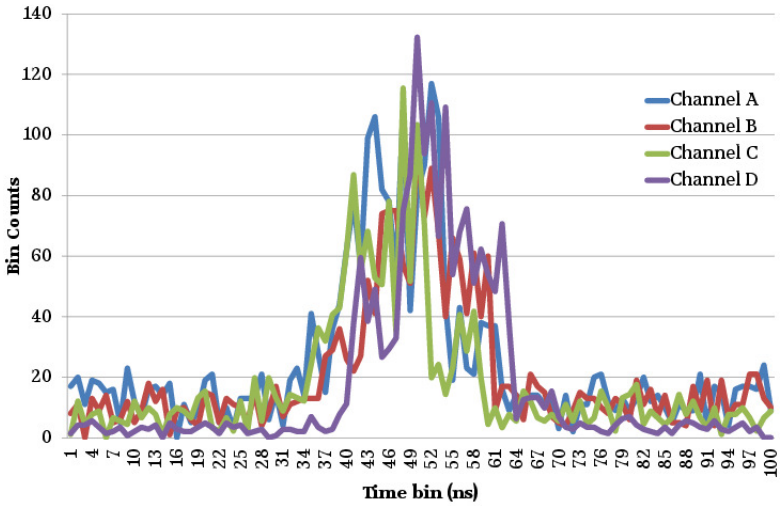


Figure 5.16: An initial estimate of the centre of each of the channel peaks is made by assuming the centre is close to the maximum bin. A constant time is added to every time tag with that channel to move this estimated maximum to the middle of the clock period.

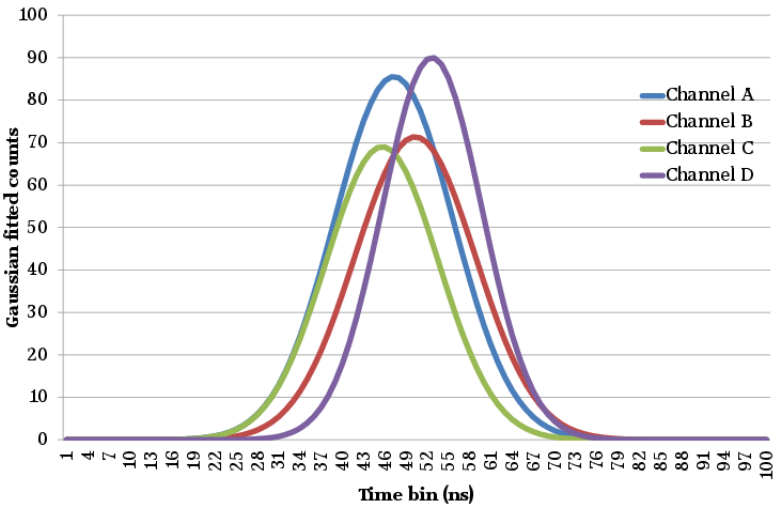


Figure 5.17: A gaussian is fitted to the peak (hence why it is moved to the middle of the clock first - if it were near the end the fitting could malfunction). A further correction is applied to move the gaussian centres to the middle of the clock.

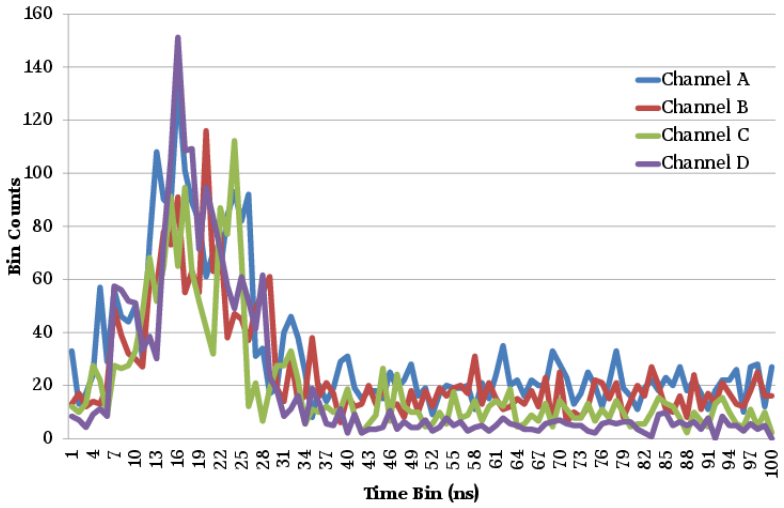


Figure 5.18: Histogram of tags divided modulo clock period once the channel synchronising offsets have been applied. This slice is taken between 3500ms-3600ms showing that the channel synchronisation does not change over the course of the transmission.

Since we are developing calibration techniques, this was used as an opportunity to apply a correction to all of the tags to account for the clock discrepancies. This process also relies on fitting gaussians to histograms of the tags divided modulo clock period, in this case knowing the original peak centre is required so after the data is offset by the rough centre position and the gaussian is fitted, the rough position value is then subtracted from the gaussian to get the original peak centre. As can be seen from a graph of histograms of subsequent 100ms slices of the tags divided modulo clock period, the peak position moves in a constant direction (figure 5.19). This is due to the clocks having a slightly different clock period, the peaks not being exactly the same distance apart is due to random fluctuations in the period and as such we will analyse every slice and obtain the average discrepancy for the transmission.

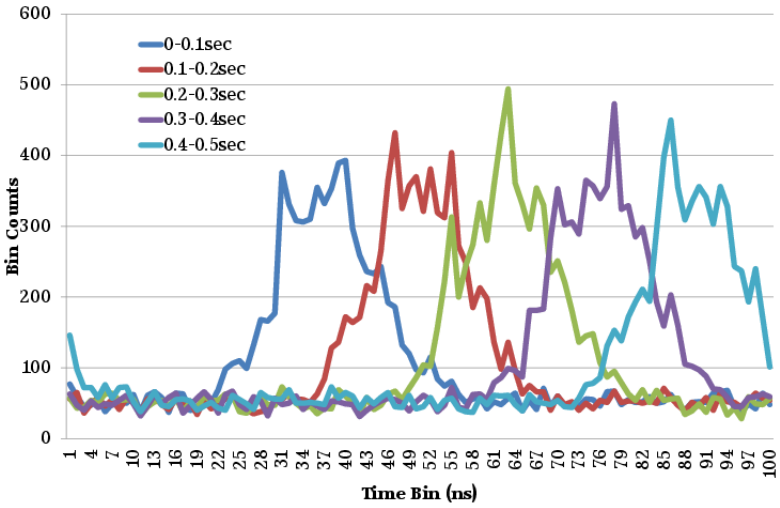


Figure 5.19: The Alice and Bob clocks are not synchronised (what Alice perceives to be 100ns and what Bob percieves to be 100ns are different), the histogram of remainders divided modulo clock period will be broadened and the peak will move when subsequent temporal slices of the tags are analysed.

For the data in figure 5.19, the average drift was found to be 131.93ns/sec. To correct this drift, every tag was taken and was transformed by:

$$NewTag = Tag + (Tag \times Drift) \tag{5.5}$$

To verify the correction, figure 5.20 shows the same slices as figure 5.19 after the correction has been applied. The entire transmission plotted in the same histogram can be seen in figure 5.21.

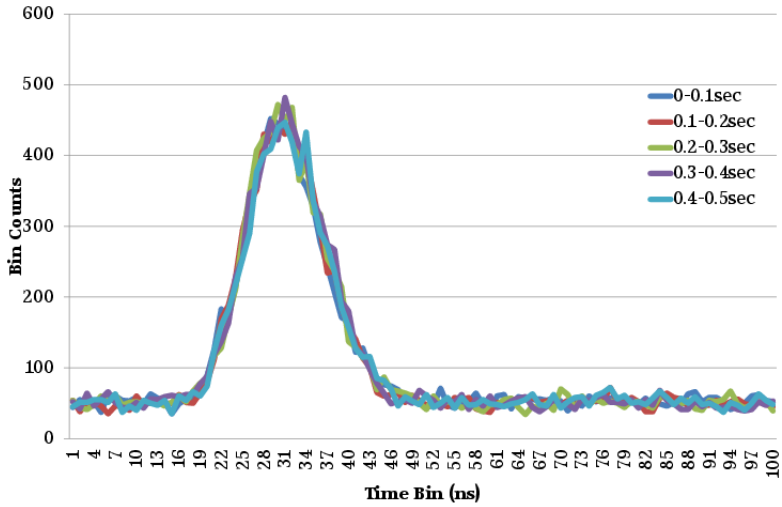


Figure 5.20: The slices from figure 5.19 corrected by determining a drift correction factor from the differences between the peaks of the slice histograms.

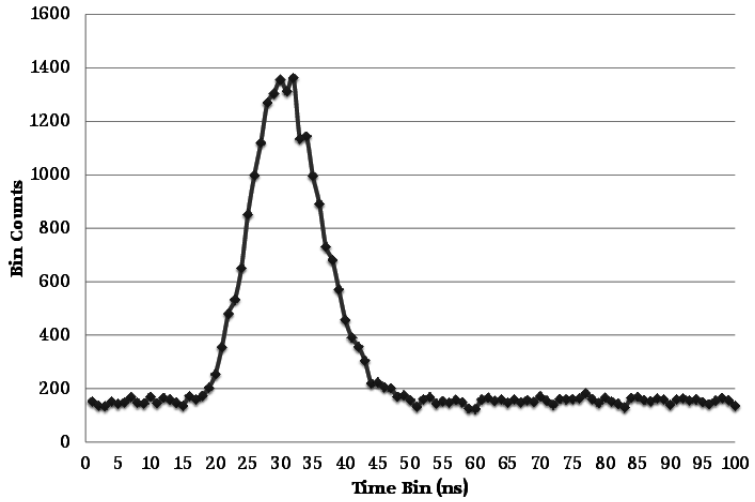


Figure 5.21: Histogram of entire transmission time tags divided modulo clock period (figure 5.20 is only the first 5 slices). Note that the width of the whole transmission corrected is narrower than that of an individual slice in figure 5.19.

With the new TIA and new calibration technique, the experiment from section 5.3 was performed again, from 15:30 on 12/11/13 to 11:00 on 13/11/13. The drop in error rate in figure 5.22 corresponds to a time around 17:30 (office lights being switched off) and the gradual increase corresponds to sunrise (07:26 - 57360 seconds from experiment start). Again the data was also plotted against the background light levels (estimated by the number of ungated tags) (figure 5.23) and was seen to be linear. There were no erroneously high data points in this data set showing that the system is suitable for long-term operation and the new TIA has solved the reliability issues.

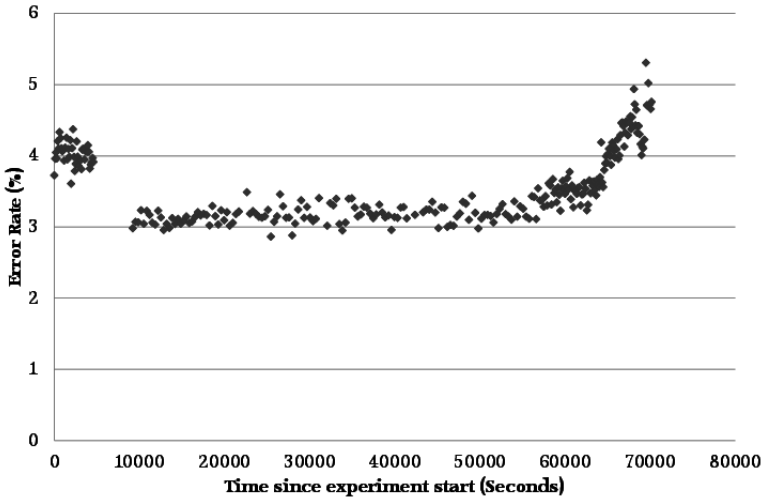


Figure 5.22: The error rate of the QKD transmission as a function of time for 19.5 hours from 15:30 12/11/13 to 11:00 13:11:13. Sharp drop around 17:30 corresponding to the office being vacated (artificial lighting switched off) and a gradual increase from sunrise (07:26 - 57360 seconds since experiment start).

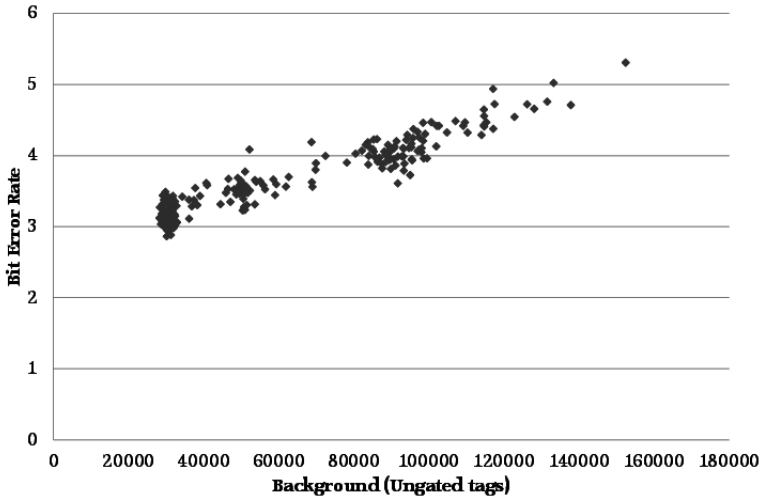


Figure 5.23: The linear relationship between error rate and background light level (estimated by number of ungated tags received).

5.4 Summary

This section has re-characterised the Bristol Low Cost Short Range QKD system and developed a modular, general purpose software suite for data analysis. This software consists of:

- Transmission isolation.
- Software clock synchronisation.
- Gating (including optimal gate width determination).
- Reconciliation.
- Error correction.

The system, however, is by no means complete, what has been achieved is a proof-of-principle "Quantum ATM" with a removable handheld Alice device. There is much to modify in the system to bring it to state of the art.

Finite Key Section 1.6.2 introduced the concept that security is not guaranteed for a certain range of parameters unless those parameters are estimated to a sufficient accuracy. This means that a certain number of signals need to be exchanged before any key can be generated at all. [137] rigorously simulated this system and concluded that the transmission time to exchange sufficient signals to generate key is unacceptably long ($\mathcal{O}(100)$) seconds with realistic background light levels).

Rate Improvements The smaller LEDs from chapter 3 and the actively quenched SPADs from chapter 4 are being designed in part to increase the possible transmission rate for the system, in the finite key regime this will reduce the time until secure key is generated back down to a practical level for a consumer system.

Protocol Change Decoy states can also provide a benefit to the key rate. It would be a trivial modification to the Alice electronics to have two driver circuits with different current limits per LED (to give the option between two different μ values) although it would add extra size. Currently the Alice LED drivers are made from discrete components but designing an integrated chip could bring this method into feasibility. Note that this method also requires 50% more random bits than regular Bennett Brassard 1984 (BB84).

Implementation Security This is not really a modification *per se* however it should always be noted that any change to the system should be considered in its resistance (or otherwise) to implementation and side channel attacks. Currently the largest issue here is that a passively quenched SPAD circuit is highly susceptible to being controlled by tuned laser pulses.

Integrate Random Bit Generation In parallel to this project, a shot noise based random bit generator is being developed. It is not yet possible to incorporate this into the system but is forthcoming.

Privacy Amplification Once error correction is performed, some form of classical privacy amplification is required. This reduces any information Eve might still have on the key to zero and is the final step in obtaining a genuinely secret key.

Chapter 6

Conclusions

6.1 Summary

In chapter 1 a background in classical cryptography was established before describing the necessity for quantum key distribution (QKD). Various QKD protocols are introduced and the Bennett Brassard 1984 (BB84) protocol was chosen as the most appropriate for this system. Thus a deeper analysis of the security of BB84 and various modifications to the protocol were discussed.

Chapter 1 then provided a brief history of free space QKD, specifically discussing the introduction of the various design choices which have been utilized in this system. Notably the techniques of beam combining for qubit generation rather than modulating a single source and the utilization of a 50:50 beam splitter in the detection optics in order to perform the random basis choice. Following this discussion, several applications for our low cost, short range system are suggested.

Chapter 2 covered the system in general, giving an idea of the components utilized and the principles governing its operation. The chapter then went on to discuss the preparations made for the SECOQC conference wherein a range of different QKD technologies were showcased and the performance of the system at that

date was presented. An average extinction ratio was measured at 6.87% and a $\approx 40,000$counts/second secret key rate (determined by the GLLP proof) was measured in indoor lighting conditions. Following this work, the finite key security proofs detailed in section 1.6 yielded a very long key expansion time, clearly insufficient for consumer applications.

In light of the finite key result from chapter 2, it was deemed necessary to increase the transmission rate. Chapter 3 and chapter 4 separated this task into improvements to the Alice and Bob devices (respectively). Work was carried out to fulfill the requirements described in table 1.1 by isolating subcomponents of the devices and carrying out research into methods of improving them especially in light of the performance measured in chapter 2.

Chapter 3 concentrated on miniaturizing the optics of quantum bit (qubit) generation for reasons of practicality and speed of operation. A new method for collimating several sources into one by pinhole diffraction was suggested and simulated. Dimensions of less than $500\mu m \times 500\mu m$ and 10-15mm length were seen as viable for the current state of the art technologies. Based on this collimation technique, a new method for producing the polarization states necessary for BB84 using microstructured gratings was introduced and tested. Whilst a polarizing effect was detected, the maximum extinction ratio was less than 7:1 which is not yet sufficient for application into the system but shows promise for the technique.

Chapter 4 concerned the efficiency and speed of the detection optics whilst increasing security. The dead time of the detectors was brought closer to the state of the art (from 3.3μs to $680ns$) at the expense of timing jitter. A modification to Active Quenching was proposed, simulated and analysed which quenches all detectors when any one detector clicks. This method was shown in simulation to not saturate in the same manner as a passive circuit where detections during another detector's dead time must be discarded. The active circuit performed acceptably in isolation

however the circuit was not stable enough for integration into the system.

An improved cooling circuit was also designed which replaced the microcontroller based circuit used previously with one based on a comparator. This circuit is simpler to set up, more efficient and more reliable.

Chapter 5 details the actual process of generating key from the components described previously. Characterisation previously carried out in chapter 2 was carried out again, most importantly the reliability of the magnetic docking mount for providing quick alignment was thoroughly tested.

A complete Python program was written to collect and analyse data to generate raw key and a basic error correction algorithm was implemented. Appendix A discusses the code structure and its performance.

A method of synchronising independent clocks by analysing the data received in temporal slices was implemented and shown to work in situations where a non sliced analysis would fail. This method also unveiled some erroneous behaviour in the time interval analyser (TIA) hardware which has since been rectified.

The error correction algorithm implemented is based on the highly interactive CASCADE algorithm. The interactivity requires many small communications between Alice and Bob, while this does not strain the connection bandwidth, the latency of the link can cause significant issues in a commercial application. Nevertheless, with correct selection of the initial block size, this algorithm performed with efficiency close to the Shannon limit - the theoretical limit for the efficiency of an error correction protocol.

In light of the finite key security analysis, discussing the secret key rate is not strictly possible with this system currently, the operation rate is far too low to exchange sufficient signals in a reasonable amount of time to generate secure key. Nevertheless, the quantum bit error rate (QBER) is still an important metric in the analysis of a QKD system and this was analysed for a variety of background

illumination rates in the lab and for an extended period of time in a naturally light room.

The QBER over a 23 hour experiment were found to track the expected background illumination (correlated with the time of day), approximately 5% during the day time, 4% in the evening and 3% at night.

6.2 Suggested Future Improvements

6.2.1 Alice

A polarization effect was measured from microstructured 1D gratings although the effect observed here was not sufficient for application into the QKD setup. Considered as a major limitation is that of the focussed ion beam (FIB) processing being unable to etch high aspect ratio features (deep and narrow trenches in this case) accurately and the relatively small area that can be patterned at once. This is not an insurmountable problem as FIB etching was chosen due to its speed and simplicity for prototyping and another technique such as e-beam lithography could be employed for better accuracy and eventually larger volume production.

The current LEDs are of large size which restricts the investigation of beam combiners such as those described in section 3.1.2. A micro LED array[1] would be a perfect test bed for further application as not only are the pitches small enough to apply the pinhole collimation techniques, the active areas are also small enough such that microfabricating polarizers across the entire surface is possible which removes the requirement for any extra focusing optics. The reduced area will hopefully lead to a lower parasitic capacitance which would allow for higher pulse rates although this is not yet a concern since the rate is currently limited by the detector dead time.

[1] `http://www.mled-ltd.com/` or `http://www.tyndall.ie/content/leds-gan` for example

Due to the commercial interest in micro LEDs, the currently available modules tend to be much larger than the 2×2 necessarily required for a QKD source (figure 6.1. This can be used positively since if the polarizers were tiled as in figure 6.2, then a small fine tuning of the alignment could be performed by testing the extinction ratios and count rates of a variety of 2×2 sub blocks in the array.

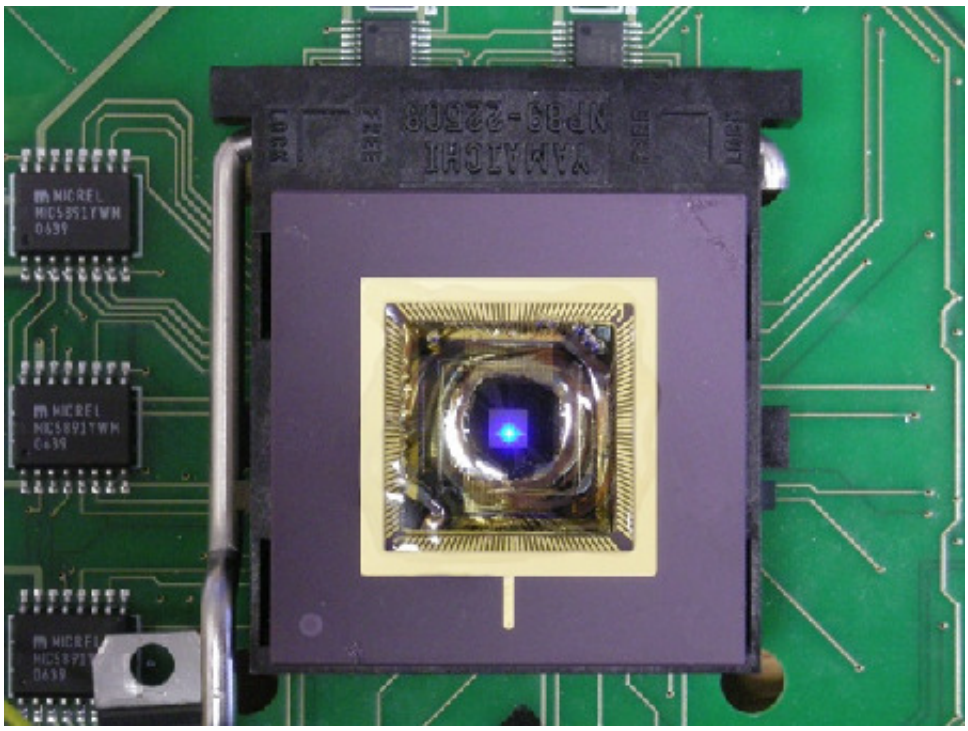


Figure 6.1: An example micro LED matrix, during fabrication, each LED in a matrix like this could have a polarizer fabricated onto it such that there is a tiled grid of QKD emitters. Image taken from `http://www.mled-ltd.com/`

Extension work eventually will also need to be carried out on further miniaturizing the drive and communication electronics in the Alice device however this is not of paramount importance since it is anticipated that the specifics of the electronics would be closely related to the final application (eg credit card, mobile phone) and therefore cannot necessarily be pre-empted.

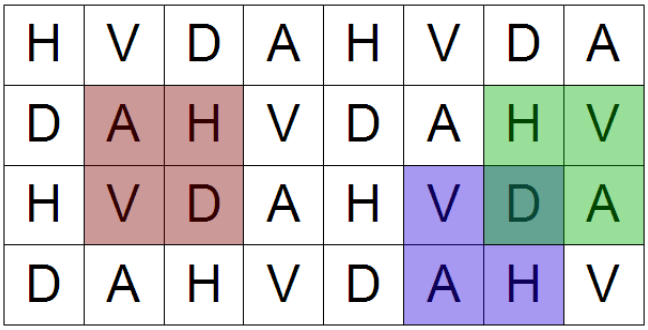


Figure 6.2: Example order of tiled polarizers allowing for fine tuning of by selecting the block with the best alignment. As shown by the coloured boxes, any 2×2 block will provide the 4 states required for BB84.

6.2.2 Bob

The $\approx 500ns$ dead time achieved in section 4.2.1 appears to approach the limit of what is available when remotely operating the single photon avalanche diodes (SPADs) using TTL components to process the signals to the quenching drivers[2]. Great success has been had by others with emitter-coupled logic (ECL) Logic however it is a technology that is largely deprecated now so investigation into other methods is necessary. Attractive technologies to pursue are complex programmable logic device (CPLD) or field programmable gate array (FPGA), both have competitive bandwidth and would allow for soft reconfiguration, a feature which would prove useful in commercial application.

The previous system [73] and all work in this thesis have utilized the C30902S-DTC avalanche photodiodes (APDs) purchased from Perkin Elmer, a performance gain may be obtained switching to a newer device; [138] suggests the Laser Components SAP500[3] as an ideal upgrade to the C30902S-DTC citing improved performance in detection efficiency (especially at 600-800nm), timing jitter, minimum achievable dead time, dark count rate and afterpulsing probability.

[2]although `http://www.potatosemi.com` appear to be making progress on producing fast TTL compatible chips

[3]`http://www.lasercomponents.com/fileadmin/user_upload/home/Datasheets/lcd/sap-series.png`

An option to increase the timing performance of the quenching circuit would be to modify the quench circuit slightly by moving the comparator which senses the avalanche signal onto the high side of the APD through a capacitor to shield the comparator from the high voltage DC bias (figure 6.3) [139, 140]. This would allow the quench voltage to be applied directly to the low side of the APD without the injecting spurious signal into the comparator (which, if unmanaged, could cause astable behaviour).

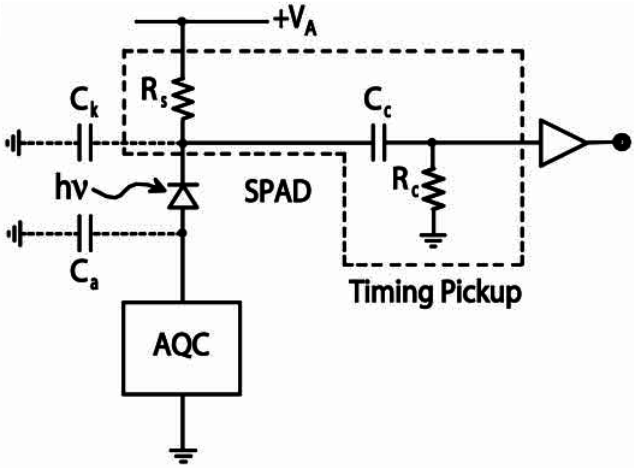


Figure 6.3: A current mode active quenching circuit with sensing performed by a capacitor on the high side of the APD. The quench is applied on the other side. This arrangement retains the simplicity of the voltage mode AQAR circuit but has much better timing performance. Image is fig 12 in [139].

Depending on the final decision on the exit pinhole of the Alice optics, a SPAD with a smaller active area than the C30902S-DTC ($500\mu m$) could be employed, this would increase performance by reducing junction capacitance (a contributor to the dead time and jitter) and there are some devices such as the id101 from idquantique[4] which have all control electronics packaged with the detector. It must be noted however that a $50\mu m$ active area would either cause high loss (due to the inability to image a spot larger than this onto it) or increased alignment difficulties.

[4] http://www.idquantique.com/scientific-instrumentation/id100-silicon-apd-single-photon-detector.html

6.2.3 Processing

Currently, to aid prototyping and development speed, the current processing code is written in Python with some targeted bottlenecks written in Cython[5]. The Numpy and Scipy libraries have provided a useful base however a migration to a faster language will be required for actual application. In fact, all processing has been demonstrated on an FPGA [141] although this solution is probably somewhat premature for this system, it is intriguing to imagine integrating the TIA and processing on the same FPGA.

To allow for independent development, the software for Alice, the TIA and the Bob processing have been developed separately, this has allowed performance analysis and bug fixing to be isolated but has meant that integrating all of these into one system has presented issues. For example, the TIA collects the time tags into RAM, saves to a file and then the analysis program then must load the tags from the file back into RAM. This is a totally unnecessary process and leads to a large delay in processing.

Another intriguing possibility regarding the software, is to involve parallel processing techniques where appropriate. An example where this technique would provide significant gains is in the synchronising step, each offset analysed is independent of the previous offsets analysed, a situation referred to as "embarrassingly parallel". This could be done either on a multi-core processor or on a general purpose graphical processing unit (GPGPU).

6.2.4 General System

Along with specific modifications to components of the system, there are also more generalised improvements which would improve the project overall. While the magnetic mount is currently sufficient for maintaining alignment, if a system were to be

[5]a compiled Python-like language which can provide performance improvements

implemented it is possible several different devices of differing specifications could be used at the same time, the best (hardware) solution to this is implementing an auto alignment system wherein a set pattern is transmitted by the Alice device and the Bob terminal measures the detections and maximises the count rate and extinction ratios by adjusting the relative xyz position, tip/tilt and rotation of the Alice and Bob optics.

With a view to being able to demonstrate this device to an audience, progress is necessary in some non scientific aspects of the project as well. Namely a clear and simple interface for interacting with the system and generating keys, preferably also including some kind of demonstration of encryption with exchanged keys as "proof" of the principle. Further to this, a modification of the Bob device to identify distinct Alice devices with a "handshake" and then store independent key pairs between different Alice devices would provide a compelling example of our suggested usage model.

Bibliography

[1] B Schneier. *Applied Cryptography.* Wiley, 1996.

[2] P W Shor. Algorithms for quantum computation: Discrete logarithms and factoring. *Proceedings, 35th Annual Symposium on Foundations of Computer Science*, pages 124–134, 1994.

[3] P Wayner. *Disappearing Cryptography: Information Hiding: Steganography and Watermarking.* Morgan Kaufmann, 2009.

[4] J Katz and Y Lindell. *Introduction to Modern Cryptography.* Chapman & Hall/CRC, 2008.

[5] A D Godley. *Herodotus: Histories Book 5, Chapter 35, Section 3.* Harvard University Press, 1920.

[6] N Biggs. *Codes: An Introduction to Information Communication and Cryptography.* Springer-Verlag, 2008.

[7] C E Shannon. Communication theory of secrecy systems. *Bell System Technical Journal*, 28(4):656715, 1949.

[8] G S Vernam. Cipher printing telegraph systems for secret wire and radio telegraphic communications. *Transactions of the American Institute of Electrical Engineers*, pages 295–301, 1926.

[9] D Stinson. *Cryptography: Theory and Practice*. CRC Press, 1995.

[10] P Ribenboim. *The Book of Prime Number Records*. Springer-Verlag, 1989.

[11] R L Rivest, A Shamir, and L Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21(2):120–126, 1978.

[12] D Boneh and R Venkatesan. Breaking RSA may be easier than factoring. *Advances in Cryptology EUROCRYPT'98*, pages 59–71, 1998.

[13] R Rivest and B Kaliski. RSA problem. *Encyclopedia of Cryptography and Security*, pages 532–536, 2003.

[14] L M K Vandersypen, M Steffen, G Breyta, C S YAnnoni, M H Sherwood, and I L Chuang. Experimental realization of Shors quantum factoring algorithm using nuclear magnetic resonance. *Nature*, 414:883–887, 2001.

[15] A Politi, J C F Matthews, and J L O'Brien. Shor's quantum factoring algorithm on a photonic chip. *Science*, 325(5945):1221, 2009.

[16] D Bernstein, J Buchmann, and E Dahmen, editors. *Post-Quantum Cryptography*. Springer-Verlag, 2009.

[17] D Dieks. Communication by EPR devices. *Physics Letters*, 92A(6):271–272, 1982.

[18] W K Wootters and W H Zurek. A single quantum cannot be cloned. *Nature*, 299:802–803, 1982.

[19] CH Bennett, F Bessette, G Brassard, L Salvail, and J Smolin. Experimental quantum cryptography. *Journal of Cryptology*, 5:3–28, 1992.

[20] S Wiesner. Conjugate coding. *ACM SIGACT News - A special issue on cryptography*, 15(1):78–88, 1983.

[21] C H Bennett and G Brassard. Quantum cryptography: Public key distribution and coin tossing. *Proceedings of IEEE International Conference on Computers Systems and Signal Processing*, pages 175–179, 1984.

[22] A K Ekert. Quantum cryptography based on bell's theorem. *Physical Review Letters*, 67(6):661–663, 1991.

[23] A Einstein, B Podolsky, and N Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 47:777–780, 1935.

[24] J Bell. On the Einstein Podolsky Rosen paradox. *Physics*, 1(3):195–200, 1964.

[25] A Ekert, J Rarity, P Tapster, and G Palma. Practical quantum cryptography based on two-photon interferometry. *Physical Review Letters*, 69(9):1293–1295, 1992.

[26] P Tapster, J Rarity, and P Owens. Violation of bell's inequality over 4 km of optical fiber. *Physical Review Letters*, 73, 1994.

[27] P Shor and J Preskill. Simple proof of security of the BB84 quantum key distribution protocol. *Physical Review Letters*, 85:441–444, 2000.

[28] M Curty, M Lewenstein, and N Lütkenhaus. Entanglement as a precondition for secure quantum key distribution. *Physical Review Letters*, 92, 2004.

[29] V Scarani and R Renner. Security bounds for quantum cryptography with finite resources. *Lecture Notes in Computer Science: Theory of Quantum Computation, Communication, and Cryptography*, 5106:83–95, 2008.

[30] R Ursin, F Tiefenbacher, T Schmitt-Manderbach, H Weier, T Scheidl, M Lindenthal, B Blauensteiner, T Jennewein, J Perdigues, P Trojek, B Ömer, M Fürst, M Meyenburg, J Rarity, Z Sodnik, C Barbieri, H Weinfurter, and A Zeilinger. Entanglement-based quantum communication over 144 km. *Nature Physics*, 3:481 – 486, 2007.

[31] C H Bennett. Quantum cryptography using any two nonorthogonal states. *Physical Review Letters*, 68(21):3121–3124, 1992.

[32] W Buttler, R Hughes, P Kwiat, S Lamoreaux, G Luther, G Morgan, J Nordholt, C Peterson, and C Simmons. Practical free-space quantum key distribution over 1 km. *Physical Review Letters*, 81(15):3283–3286, 1998.

[33] D Bruß. Optimal eavesdropping in quantum cryptography with six states. *Physical Review Letters*, 81(14):3018–3021, 1998.

[34] "Event-Ready-Detectors" Bell Experiment via Entanglement Swapping. M zukowski and a zeilinger and m a horne and a k ekert. *Physical Review Letters*, 71:4287–4290, 1993.

[35] D Collins, N Gisin, and H De Reidmatten. Quantum relays for long distance quantum cryptography. *Journal of Modern Optics*, 52:735–753, 2007.

[36] C H Bennett, G Brassard, S Popescu, B Schumacher, J Smolin, and W Wootters. Purification of noisy entanglement and faithful teleportation via noisy channels. *Physical Review Letters*, 76:722–725, 1996.

[37] N Sangouard, C Simon, H De Riedmatten, and N Gisin. Quantum repeaters based on atomic ensembles and linear optics. *Review of Modern Optics*, 83:33–80, 2011.

[38] L-M Duan, M Lukin, J I Cirac, and P Zoller. Long-distance quantum communication with atomic ensembles and linear optics. *Nature*, 414:413–418, 2001.

[39] C Kurtsiefer, P Zarda, S Mayer, and H Weinfurter. The breakdown flash of silicon avalanche photodiodes-back door for eavesdropper attacks? *Journal of Modern Optics*, 48(13):2039–2047, 2001.

[40] S Nauerth, M Fürst, T Schmitt-Manderbach, H Weier, and H Weinfurter. Information leakage via side channels in freespace BB84 quantum cryptography. *New Journal of Physics*, 11, 2008.

[41] H Weier, H Krauss, M Rau, M Fürst, S Nauerth, and H Weinfurter. Quantum eavesdropping without interception: an attack exploiting the dead time of single-photon detectors. *New Journal of Physics*, 13:073024, 2011.

[42] J Friedman. Tempest: A signal problem. *NSA Cryptologic Spectrum*, 1972.

[43] V Scarani and C Kurtsiefer. The black paper of quantum cryptography: real implementation problems. *arXiv:0906.4547v1*, 2009.

[44] M Eisaman, J Fan, A Migdall, and S Polyakov. Invited review article: Single-photon sources and detectors. *Review Of Scientific Instruments*, 82(7):071101, 2011.

[45] Y Hu, X Peng, T Li, and H Guo. On the Poisson approximation to photon distribution for faint lasers. *Physics Letters A*, 367(3):173–176, 2007.

[46] W-Y Hwang. Quantum key distribution with high loss: Toward global secure communication. *Physical Review Letters*, 91(5):057901, 2003.

[47] A Lacaita, F Zappa, S Bigliardi, and M Manfredi. On the bremsstrahlung origin of hot-carrier-induced photons in silicon devices. *IEEE Transactions On Electron Devices*, 40(3):577–582, 1993.

[48] V Makarov. Exploiting the saturation mode of passively quenched avalanche photodiodes to attack quantum cryptosystems. *Proceedings of the Optical Society of Korea Annual Meeting 2008*, pages 417–418, 2008.

[49] V Makarov. Controlling passively-quenched single photon detectors by bright light. *New Journal of Physics*, 11:065003, 2009.

[50] Z L Yuan, J F Dynes, and A J Shields. Avoiding the blinding attack in QKD. *Nature Photonics*, 4:800801, 2010.

[51] I Gerhardt, Q Liu, A Lamas-Linares, J Skaar, C Kurtsiefer, and V Makarov. Full-field implementation of a perfect eavesdropper on a quantum cryptography system. *Nature Communications*, 2:349, 2011.

[52] L Lydersen and V MakarovAND J Skaar. Secure gated detection scheme for quantum cryptography. *Physical Review A*, 83:032306, 2011.

[53] M Ardehali, H F Chau, and H-K Lo. Efficient quantum key distribution. *arXiv:quant-ph/9803007*, 1999.

[54] H-K Lo, H F Chau, and M Ardehali. Efficient quantum key distribution scheme and a proof of its unconditional security. *Journal of Cryptology*, 18:133165, 2005.

[55] M Curty, T Moroder, X Ma, H-K Lo, and N Lütkenhaus. Upper bounds for the secure key rate of decoy state quantum key distribution. *Physical Review A*, 79(3):032335, 2009.

[56] M Curty, X Ma, B Qi, and T Moroder. Passive decoy state quantum key distribution with practical light sources. *Physical Review A*, 81(2):022310, 2009.

[57] V Scarani, A Acín, G Ribordy, and N Gisin. Quantum cryptography protocols robust against photon number splitting attacks for weak laser pulse implementations. *Physical Review Letters*, 92(5):057901, 2004.

[58] C Branciard, N Gisin, B Kraus, and V Scarani. Security of two quantum cryptography protocols using the same four qubit states. *Physical Review A*, 72:032301, 2005.

[59] D Mayers. Unconditional security in quantum cryptography. *Journal of the ACM*, 48(3):351–406, 2001.

[60] H Inamori, N Lütkenhaus, and D Mayers. Unconditional security of practical quantum key distribution. *The European Physical Journal D*, 41(3):599–627, 2001.

[61] M Koashi and J Preskill. Secure quantum key distribution with an uncharacterized source. *Physical Review Letters*, 90(5):057902, 2002.

[62] D Gottesman, H-K Lo, N Lütkenhaus, and J Preskill. Security of quantum key distribution with imperfect devices. *Quantum Information and Computation*, 4(5):325–360, 2004.

[63] X Ma. Unconditional security at a low cost. *Physical Review A*, 74(5):052325, 2006.

[64] I Devetak and A Winter. Distillation of secret key and entanglement from quantum states. *Proceedings of the Royal Society A*, 461:207235, 2005.

[65] V Scarani, H Bechmann-Pasquinucci, N J Cerf, M Dušek, N Lütkenhaus, and M Peev. The security of practical quantum key distribution. *Review of Modern Physics*, 81:13011350, 2009.

[66] V Scarani and R Renner. Quantum cryptography with finite resources: Unconditional security bound for discrete-variable protocols with one-way postprocessing. *Physical Review Letters*, 100:200501, 2008.

[67] R Y Q Cai and V Scarani. Finite-key analysis for practical implementations of quantum key distribution. *New Journal of Physics*, 11:045024, 2008.

[68] J G Rarity, P C M Owens, and P R Tapster. Quantum random-number generation and key sharing. *Journal of Modern Optics*, 41, 1994.

[69] J G Rarity, P R Tapster, and P M Gorman. Secure free-space key exchange to 1.9km and beyond. *Journal of Modern Optics*, 48:13:1887–1901, 2001.

[70] R J Hughes, J E Nordholt, D Derkacs, and C G Peterson. Practical free-space quantum key distribution over 10 km in daylight and at night. *New Journal of Physics*, 4:43.143.14, 2002.

[71] C Kurtsiefer, P Zarda, M Halder, H Weinfurter, P M Gorman, P R Tapster, and J G Rarity. A step towards global key distribution. *Nature*, 419:450, 2002.

[72] J L Duligall, M S Godfrey, K A Harrison, W J Munro, and J G Rarity. Low cost and compact quantum key distribution. *New Journal of Physics*, 8:249, 2006.

[73] J Duligall. *Compact and Low Cost Quantum Cryptography*. PhD thesis, 2007.

[74] http://www.idquantique.com.

[75] http://www.magiqtech.com.

[76] http://www.quintessencelabs.com/.

[77] EMV integrated circuit card specifications for payment systems book 2: Security and key management version 4.2, 2008.

[78] Barclays PINSentry user guide - item ref 9907259 - (04-2007), 2007.

[79] R Nock, N Dahnoun, and J Rarity. Low cost timing interval analyzers for quantum key distribution. *Instrumentation and Measurement Technology Conference (I2MTC), 2011 IEEE*, pages 475–479, 2011.

[80] J L O'Brien, A Furusawa, and J Vučković. Photonic quantum technologies. *Nature Photonics*, 3:687–695, 2009.

[81] P Michler, A Kiraz, C Becher, W Schoenfeld, P Petroff, L Zhang, E Hu, and A Imamoğlu. A quantum dot single-photon turnstile device. *Science*, 290:2282–2285, 2000.

[82] J Rarity, P Tapster, and E Jakeman. Observation of sub-poissonian light in parametric down-conversion. *Optics Communications*, 62:201–206, 1987.

[83] C Hong and L Mandel. Experimental realization of a localized one-photon state. *Physical Review Letters*, 56(1):58–60, 1986.

[84] T Gaebel, I Popa, A Gruber, M Domhan, F Jelezko, and J Wrachtrup. Stable single-photon source in the near infrared. *New Journal of Physics*, 6:98, 2004.

[85] E Wooten, K Kissa, A Yi-Yan, E Murphy, D Lafaw, P Hallemeier, D Maack, D Attanasio, D Fritz, G McBrien, and D Bossi. A review of lithium niobate modulators for fiber-optic communications systems. *IEEE Journal Of Selected Topics In Quantum Electronics*, 6(1):69–82, 2000.

[86] E F Schubert, N E J Hunt, R J Malik, M Micovic, and D L Miller. Temperature and modulatilon characteristics of resonant-cavity light-emitting diodes. *Journal of Lightwave Technology*, 14:1721, 1996.

[87] E Hecht. *Optics*. Addison-Wesley, 1998.

[88] M L Boas. *Mathematical Methods in the Physical Sciences*. Wiley, 1984.

[89] H-K Lo and J Preskill. Security of quantum key distribution using weak coherent states with nonrandom phases. *Quantum Information and Computation*, 7:431–458, 2007.

[90] R S Longhurst. *Geometrical and Physical Optics*. Longman, 3rd edition, 1973.

[91] J Meyer-Arendt. *Introduction to Classical and Modern Optics*. Prentice-Hall, 1989.

[92] Perkin Elmer. Biomedical avalanche photodiodes C30902SH, C30921SH, C30902SH-TC, C30902SH-DTC, 2007.

[93] R G W Brown, K D Ridley, and J G Rarity. Characterization of silicon avalanche photodiodes for photon correlation measurements. 1: Passive quenching. *Applied Optics*, 25(22):4122–4126, 1986.

[94] Perkin Elmer. Avalanche photodiode: A user guide. understanding avalanche photodiode for improving system performance, 2006.

[95] A Spinelli and L Lacaita. Physics and numerical simulation of single photon avalanche diodes. *IEEE Transactions On Electron Devices*, 44(11), 1997.

[96] M A Itzler, R Ben-Michael, C F Hsu, K Slomkowski, A Tosi, S Cova, F Zappa, and R Ispasoiu. Single photon avalanche diodes (SPADs) for 1.5 μm photon counting applications. *Journal of Modern Optics*, 54:2-3, 2007.

[97] S Cova, M Ghioni, A Lacaita, C Samori, and F Zappa. Avalanche photodiodes and quenching circuits for single photon detection. *Applied Optics*, 35(12):1956–1976, 1996.

[98] A M Lynch. *Low cost and flexible electronics for quantum key distribution and quantum information*. PhD thesis, 2010.

[99] M Peev, C Pacher, R Alléaume, C Barreiro, J Bouda, W Boxleitner, T Debuisschert, E Diamanti, M Dianati, J F Dynes, S Fasel, S Fossier, M Fürst, J-D Gautier, O Gay, N Gisin, P Grangier, A Happe, Y Hasani, M Hentschel, H Hübel, G Humer, T Länger, M Legré, R Lieger, J Lodewyck, T Lornser, N Lütkenhaus, A Marhold, T Matyus, O Maurhart, L Monat, S Nauerth, J-B Page, A Poppe, E Querasser, G Ribordy, S Robyr, L Salvail, A W Sharpe, A J Shields, D Stucki, M Suda, C Tamas, T Themel, R T Thew, Y Thoma, A Treiber, P Trinkler, R Tualle-Brouri, F Vannel, N Walenta, H Weier, H Weinfurter, I Wimberger, Z L Yuan, H Zbinden, and A Zeilinger. The SECOQC quantum key distribution network in vienna. *New Journal of Physics*, 11:075001, 2009.

[100] D Stucki, N Brunner, N Gisin, V Scarani, and H Zbinden. Fast and simple one-way quantum key distribution. *Applied Physics Letters*, 87:194108, 2008.

[101] J Dynes, Z Yuan, A Sharpe, and A Shields. Practical quantum key distribution over 60 hours at an optical fiber distance of 20km using weak and vacuum decoy pulses for enhanced security. *Optics Express*, 15(13):8465–8471, 2007.

[102] A Treiber, A Poppe, M Hentschel, D Ferrini, T Lorünser, E Querasser, T Matyus, H Hübel, and A Zeilinger. Fully automated entanglement-based quantum cryptography system for telecom fiber networks. *New Journal of Physics*, 11:045013, 2009.

[103] J Lodewyck, M Bloch, R García-Patrón, S Fossier, E Karpov, E Diamanti, T Debuisschert, N Cerf, R Tualle-Brouri, S McLaughlin, and P Grangier. Quantum key distribution over 25 km with an all-fiber continuous-variable system. *Physical Review A*, 76:042305, 2007.

[104] H Weier, T Schmitt-Manderbach, N Regner, C Kurtsiefer, and H Weinfurter. Free space quantum key distribution: Towards a real life application. *Fortschritte der Physik*, 54(8-10):840–845, 2006.

[105] T Länger and G Lenhart. Standardization of quantum key distribution and the ETSI standardization initiative ISG-QKD. *New Journal of Physics*, 11:055051, 2009.

[106] Newport optics product catalog 2008/9 page 730.

[107] P Reddy, H Sharma, and D Paulraj. Multi channel wi-fi sniffer. *WiCOM '08. 4th International Conference on Wireless Communications, Networking and Mobile Computing*, pages 1–6, 2008.

[108] http://www.wireshark.org/.

[109] K Schmid, E Frins, H Schmitzer, and W Dultz. Beam mixing with a pinhole. *Journal of the Optical Society of America A*, 22(12):2672–2676, 2005.

[110] Agilent Technologies. Subminiature high performance AlInGaP LED lamps, 2001.

[111] E Land. Light-polarizing body. *United States Patent 2165973*, 1939.

[112] M Bloemer and J Haus. Versatile waveguide polarizer incorporating an ultrathin discontinuous silver film. *Applied Physics Letters*, 61(14):1619–1621, 1992.

[113] G Bird and M Parrish. The wire grid as a near-infrared polarizer. *Journal Of The Optical Society Of America*, 50(9):886–891, 1960.

[114] J B Young, H A Graham, and E W Peterson. Wire grid infrared polarizer. *Applied Optics*, Vol. 4, No. 8:1023–1026, 1965.

[115] J J Wang, F Walters, X Liu, P Sciortino, and X Deng. High-performance, large area, deep ultraviolet to infrared polarizers based on 40 nm line/78 nm space nanowire grids. *Applied Physics Letters*, 90, 2007.

[116] S-W Ahn, K-D Lee, J-S Kim S H Kim, J-D Park, S-H Lee, and P-W Yoon. Fabrication of a 50nm half-pitch wire grid polarizer using nanoimprint lithography. *Nanotechnology*, 16:1874–1877, 2005.

[117] Z Y Yang and Y F Lu. Broadband nanowire-grid polarizers in ultraviolet-visible-near-infrared regions. *Optics Express*, 15(15):9510–9519, 2007.

[118] M Young. *Optics and Lasers*. Springer-Verlag, 1984.

[119] M Born and E Wolf. *Principles of Optics*. Pergammon Press, 1964.

[120] D Lide, editor. *CRC Handbook of Chemistry and Physics 78th Edition*. CRC Press, 1997.

[121] T Ebbesen, H Lezec, H Ghaemi, T Thio, and P Wolff. Extraordinary optical transmission through sub-wavelength hole arrays. *Nature*, 391:667–669, 1998.

[122] U Schroter and D Heitmann. Surface-plasmon-enhanced transmission through metallic gratings. *Physical Review B*, 58(23):419–421, 1998.

[123] T Thio, K Pellerin, R Linke, H Lezec, and T Ebbesen. Enhanced light transmission through a single subwavelength aperture. *Optics Letters*, 26(24):1972–1974, 2001.

[124] F Baida, D Van Labeke, and B Guizal. Enhanced confined light transmission by single subwavelength apertures in metallic films. *Applied Optics*, 42(34):6811–6815, 2003.

[125] L Chen. Optical transmission spectroscopy measurements of plasmonic nanoantennas and gratings. Master's thesis, University of Bristol, 2011.

[126] S M Sze. *Physics of Semiconductor Devices*. Wiley, 2nd edition, 1981.

[127] T P Lee. Effect of junction capacitance on the rise time of LED's and on the turn-on delay of injection lasers. *The Bell System Technical Journal*, 54:53–68, 1975.

[128] S Park, E Sim, J-W Park, J-S Sim, H-W Song, S-H Oh, and Y Baek. Temperature, current, and voltage dependences of junction failure in PIN photodiodes. *ETRI Journal*, 28(5):555–560, 2006.

[129] D J Rogers, J C Bienfang, A Nakassis, H Xu, and C W Clark. Detector deadtime effects and paralyzability in high-speed quantum key distribution. *New Journal of Physics*, 9, 2007.

[130] V Burenkov, B Qi, B Fortescue, and H-K Lo. Security of high speed quantum key distribution with finite detector dead time. *arXiv:1005.0272v1*, 2010.

[131] H Xu, L Ma, J C Bienfang, and X Tang. Influence of avalanche-photodiode dead time on the security of high-speed quantum-key distribution systems. *Conference on Lasers and Electro-Optics, 2006 and 2006 Quantum Electronics and Laser Science Conference. CLEO/QELS 2006*, 2006.

[132] H Dautet, P Deschamps, B Dion, A MacGregor, D MacSween, R McIntyre, C Trottier, and P Webb. Photon counting techniques with silicon avalanche photodiodes. *Applied Optics*, 32:3894–3900, 1993.

[133] P Horowitz and W Hill. *The Art of Electronics*. Cambridge University Press, 1980.

[134] T D McGee. *Principles and Methods of Temperature Measurement*. Wiley, 1988.

[135] N Gisin, G Ribordy, W Tittel, and H Zbinden. Quantum cryptography. *Reviews of Modern Physics*, 74:145–195, 2002.

[136] M Dušek, N. Lütkenhaus, and M Hendrych. Quantum cryptography. *Progress in Optics*, 49:381–454, 2006.

[137] M S Godfrey. *Reconciliation and estimation for a short range quantum cryptography system*. PhD thesis, 2010.

[138] M Stipĉević, H Skenderović, and D Gracin. Characterization of a novel avalanche photodiode for single photon detection in VIS-NIR range. *Optics Express*, 18, 2010.

[139] S Cova, M Ghioni, A Lotito, I Rech, and F Zappa. Evolution and prospects for single-photon avalanche diodes and quenching circuits. *Journal of Modern Optics*, 51, 2004.

[140] S Cova, M Ghioni, and F Zappa. Circuit for high precision detection of the time of arrival of photons falling on single photon avalanche diodes. *European Patent EP20010200851*, 2001.

[141] H-F Zhang, J Wang, K Cui, C-L Luo, S-Z Lin, L Zhou, H Liang, T-Y Chen, K Chen, and J-W Pan. A real-time QKD system based on FPGA. *Journal of Lightwave Technology*, 30:3226, 2012.

BIBLIOGRAPHY

Appendix A

QKD Program Performance

A.1 Initial Code

The QKD processing task was split into sections corresponding to its basic functionality and each was timed separately. The sections are as follows:

Bob IO Since the program is currently standalone, it loads a file consisting of [CHANNEL] [COUNTER] and [PHASE], where [CHANNEL] is the time interval analyser (TIA) channel that triggered (1, 2, 4, 8 for singles, any sum of these for coincidences) and [COUNTER] and [PHASE] are the coarse clock and delay tap which form the time tag. These are processed into time tags by multiplying the [COUNTER] value by the clock rate and looking up the appropriate [PHASE] value from an automatically (by the TIA) generated calibration histogram.

Tag Sorting The tags are then checked for integrity and anomalies rejected. Possible anomalies include: zero channel, non sequential tags, abnormally large tags. This is also the section which deals with randomly assigning the coincidence counts to one of their constituent channels.

Coarse/Fine Boundary Finding These are discussed in detail in section 5.2.1. In brief the intervals between adjacent tags are analysed, a decrease in the average interval signifies the transmission window. The coarse pass concerns itself with every nth tag for speed, the fine pass focuses on the points flagged by the coarse search and finds more exact values.

Gating Time The data is divided modulo the transmission clock and the remainders are plotted on a histogram. All remainders outside of a window (the gate) of the maximum of the histogram are considered background. The integer part of the modulo division is then considered as the tag number and the timing data is thrown away.

Alice IO The subset of the transmitted data required for synchronisation and error estimation is loaded.

Sync Finding The transmitted string is compared to the gated tags for offsets spreading from zero until a greater than 85% match is found. If this step fails all the channels can be mapped to a different permutation in case Alice and Bob's channel names do not correspond to each other.

Key assembly The offset determined in the sync set is applied to one of the data sets. For the sake of characterisation, the error rate of the full data set is also obtained here (obviously this would not occur for *actual* key generation).]

Error Correction The error correction algorithm described in section 5.2.4 is performed. Again, the error rate of the full data set can be calculated here for characterisation.

Some sample key exchange timings are displayed in table A.1.

Power Meter	0.05	0.08	0.1	0.15	0.22	0.24	0.48	2.5*	4.3*
Ungated Tags	371621	405718	471399	496433	493491	536978	709440	387517	394875
Bob IO time	7.34	8.72	12.03	16.03	19.53	14.36	28.78	7.58	8.58
Tag Sorting Time	1.26	1.14	1.44	1.74	2.04	1.62	2.36	0.99	1.08
Coarse Boundary Finding	4.70	3.98	5.64	7.46	9.15	6.71	11.57	3.36	4.29
Fine Boundary Finding	1.50	1.18	1.23	1.25	1.29	1.08	1.32	0.96	1.46
Gating Time	4.14	3.61	4.31	4.48	4.40	4.89	6.45	4.21	4.31
Alice IO time	2.83	2.55	2.45	2.46	2.48	2.49	2.82	2.38	2.69
Sync Finding Time	332.41	143.48	171.35	87.45	62.50	125.78	8.41	465.66	260.01
(Sync Offset Value)	(3250)	(1349)	(1639)	(790)	(560)	(1150)	(19)	(3980)	(2273)
Time assembling keys	28.14	10.51	10.92	11.23	10.82	11.79	11.11	11.92	11.79
Error Correction Time	3.64	6.92	4.44	4.82	4.57	5.55	6.74	4.29	3.70
Running Time (mins)	6.43	3.04	3.56	2.28	1.95	2.90	1.33	8.36	4.96

Table A.1: Times (in seconds unless noted) for various processes in the QKD processing code. The "Power Meter" values are from a power meter adjacent to the input lens of Bob, this correlates with the background and was used as an identifier for the data files. The background was generated by a small red LED adjacent to the Alice optics; for the values marked with "*", a white light was used.

Some trends apparent in this data:

- Most of the times increase for increasing background. This is not surprising since a set *duration* is measured and since the signal is roughly constant, more background = more tags.

- While the coarse boundary finding is dependent on background, the fine boundary finding isn't. This is because the coarse method has to process all of the tags however the fine method only processes the output of the coarse method which is always the same.

- Sync time is unrelated to background. This is surprising since one would expect that the higher background might obscure the exact start point of the transmission. Obviously the sync time and the offset are directly related.

A.2 Numpy optimizations

Between initial submission of this work an examination, the entire code was rewritten using Numpy, a suite of optimized numerical functions for Python written in C. As can be seen in table A.2, this has had a varying degree of success depending on the task. This is because some of the required functions perform exactly the task required (eg. for smoothing the intervals there exists a convolution function which performs this efficiently).

In order to accommodate the functions, the order of processing was changed slightly so a description of the steps is below:

Data loading Python was found to not be fast enough to aquire the data from the TIA thus an external library needs to be wrapped to integrate into the program. For now the data files are loaded from disk.

Tag Processing The 4 independent TIA channels do not guarantee that the tags are in chronological order in the list, thus the sorting described previously is still necessary. This step can also filter tags out based on their channel (if required for testing).

Isolate Transmission Same as the boundary finding above. [Slice, Filter] Tags falling within one detector dead time of the previous tag are removed (for security), the time tags in the transmission are shifted such that the first tag is zero (which makes synchronizing the data easier since the indices will start near 0) and the data is split into slices of a set duration

Correct Clock Drift The process described in section 5.3.1 is performed, correcting the drifts due to unsynchronised clocks in the Alice and Bob devices and differences in the TIA channels.

Find Data Sync Offset Same as above, even though the drift is corrected, the analysis is still performed in slices which allows for the first slice synchronising offset to be used as an initial guess for the next slice. This is usually the correct guess so the remaining slice offsets can be determined without searching a large range. As in table A.1, the Synchronising offset is shown in brackets beneath this number. The sync finding time is linearly proportional to the synchronising offset.

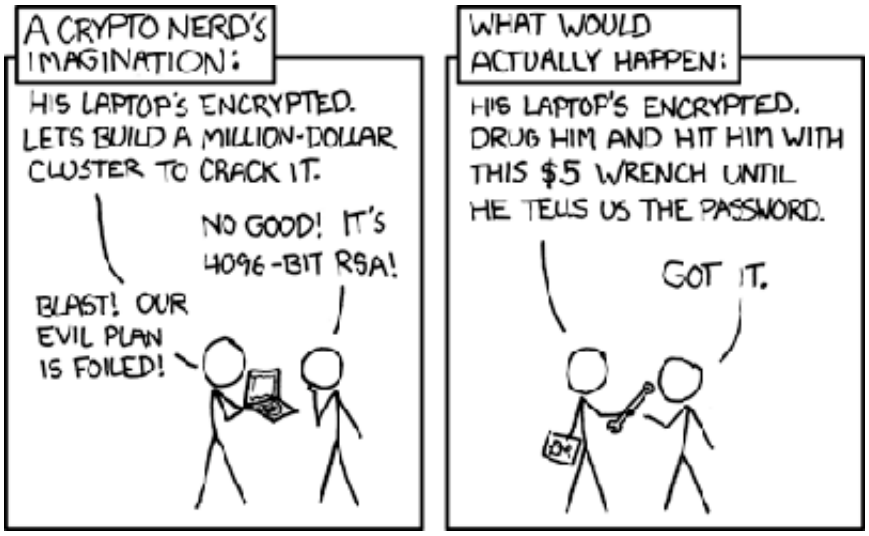
Total Time To give an impression of the processing time, the total running time is displayed along with the processing time, ie, without the data loading.

									Average
Alice data loading	13.10	13.68	13.09	13.08	13.08	13.09	13.15	13.09	13.17
Bob data loading	14.36	15.04	15.25	15.03	15.61	15.04	14.85	14.37	14.94
Tag processing	4.59	4.49	4.35	4.44	4.35	4.28	4.10	4.34	4.37
Isolate transmission	0.05	0.05	0.04	0.04	0.07	0.04	0.04	0.04	0.05
Slice, Filter	1.54	1.55	1.18	1.22	1.59	1.29	1.29	1.17	1.35
Correct clock drift	7.05	7.39	6.35	6.70	7.43	6.65	6.54	6.41	6.81
Gate data	3.26	3.99	3.12	3.38	3.36	3.22	3.21	3.08	3.33
Find data sync offset	17.50	39.47	21.11	41.44	32.96	11.11	32.88	41.71	29.77
(Best Offset)	342	850	448	1022	707	235	746	959	N/A
Total Time	61.45	85.66	64.49	85.35	78.45	54.72	76.07	84.21	73.80
Total (Excluding loading)	33.99	56.95	36.16	57.23	49.77	26.60	48.06	56.75	45.69

Table A.2: Processing time of 8 collected data sets for the rewritten QKD analysis program

Some observations:

- The tags take longer to load because the new TIA outputs the tags as floats rather than the "Counter" and "Phase" integers which need processing, these floats are then loaded into a Numpy array which takes longer but the array can be processed much more efficiently.
- The processing is also longer as the independent TIA channels produce more non-sequential time tags (theoretically, the previous one should not have produced any).
- The numpy.diff() and numpy.convolve() functions almost entirely perform the boundary finding process and as such a 20× speed increase is observed.
- A contributor to the sync time being shorter is the use of Numpy arrays however there does not exist a standard function for this process so the speed up is not as marked as for the transmission finding. This only contributes to roughly halving the processing time per offset, the other increase is that the transmission finding algorithm being so much faster allows a higher resolution search which can determine the transitions more accurately (and as such their indices will be closer to zero).
- Error correction was not performed on these data sets as the algorithm has not changed at all.
- The data loading now accounts for nearly half of the processing time.



Comic from `http://xkcd.com/538/`